

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 10:47:27 ; Search time 1415.77 Seconds

(without alignments)

10489.161 Million cell updates/sec

Title: US-10-017-372E-14

Perfect score: 363

Sequence: 1 gccctggataccaactactg.....gttcctgcaagtgcagctga 363

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

```
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

```
1:  gb ba:*
```

```
2: gb_htg:*
```

```
3: gb in:*
```

4: gb om: *

```
5: gb_ov:*
```

```
6: gb_pat:*
```

7: gb ph:*

8: gb pl:*

9: gb_pr:*

10: gb ro:*

```
11:  gb sts:*
```

```
12:  gb_sy:
```

```
13:  gb_un:*
```

```
14:  gb_vi:*
```

15: em ba:*

```
16:  em_fun:*
```

```
17:  em_hum:*
```

```
18:  em_in:*
```

```
19:  em_mu:*
```

20: em_om: *

21: em_or:*

```
22:  em_ov:*
```

```
23:  em_pat:*
```

24: em_ph:*

```
25:  em_pl:*
```

26: em ro:*

27: em sts:*

```

28:  em_un:*
29:  em_vi:*
30:  em_htg_hum:*
31:  em_htg_inv:*
32:  em_htg_other:*
33:  em_htg_mus:*
34:  em_htg_pln:*
35:  em_htg_rod:*
36:  em_htg_mam:*
37:  em_htg_vrt:*
38:  em_sy:*
39:  em_htgo_hum:*
40:  em_htgo_mus:*
41:  em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	306.8	84.5	1326	6	AX338213	AX338213 Sequence	
2	306.8	84.5	1605	4	SSTGFBR	Y00111 Porcine mRN	
3	306.8	84.5	1750	4	GGTGFB1	X12373 Porcine mRN	
4	306.8	84.5	2221	4	AF461808	AF461808 Sus scrof	
5	306.8	84.5	3206	4	PIGTGFB1A	M23703 Sus scrofa	
6	281.2	77.5	339	6	AR036686	AR036686 Sequence	
7	281.2	77.5	489	6	AX455100	AX455100 Sequence	
8	281.2	77.5	1173	4	OATGFB1	X76916 O.aries mRN	
9	281.2	77.5	1369	4	DOGTGFB1A	L34956 Canine tran	
10	281.2	77.5	1561	9	AGMTGFB	M16658 Simian tran	
11	279.6	77.0	339	6	A18277	A18277 H.sapiens T	
12	279.6	77.0	339	6	A23751	A23751 TGF-betal c	
13	279.6	77.0	339	6	A48549	A48549 Sequence 1	
14	279.6	77.0	339	6	A48563	A48563 Sequence 1	
15	279.6	77.0	339	6	I56854	I56854 Sequence 1	
16	279.6	77.0	650	6	AX336646	AX336646 Sequence	
17	279.6	77.0	650	9	HUMTGFB1A	M38449 Homo sapien	
18	279.6	77.0	699	6	I05434	I05434 Sequence 4	
19	279.6	77.0	862	6	I03312	I03312 Sequence 3	
20	279.6	77.0	1176	6	AX481432	AX481432 Sequence	
21	279.6	77.0	1176	6	AX615127	AX615127 Sequence	
22	279.6	77.0	1176	6	AX615128	AX615128 Sequence	
23	279.6	77.0	1560	6	I08268	I08268 Sequence 2	
24	279.6	77.0	1746	9	BC022242	BC022242 Homo sapi	
25	279.6	77.0	1780	9	BC000125	BC000125 Homo sapi	
26	279.6	77.0	1780	9	BC001180	BC001180 Homo sapi	
27	279.6	77.0	1821	6	E03028	E03028 DNA encodin	
28	279.6	77.0	2537	6	A06669	A06669 Synthetic m	
29	279.6	77.0	2745	9	HSTGFB1	X02812 Human mRNA	
30	279.2	76.9	337	6	AR198575	AR198575 Sequence	
31	278.6	76.7	1560	6	I06216	I06216 Sequence 2	
32	278	76.6	339	4	OCAF000133	AF000133 Oryctolag	
33	278	76.6	1117	4	BOVTGFB	M36271 Bovine tran	

34	278	76.6	1266	6	AX455875	AX455875 Sequence
35	278	76.6	2527	6	E00973	E00973 cDNA encodi
36	277.6	76.5	1173	9	BT007245	BT007245 Homo sapi
37	277.6	76.5	1173	12	BT007866	BT007866 Synthetic
38	277	76.3	334	4	SHPGFB1W	L36038 Ovis aries
39	277	76.3	1569	6	I06221	I06221 Sequence 3
40	275.2	75.8	1173	4	ECRGFB1	X99438 E.caballus
41	275.2	75.8	1187	4	AF175709	AF175709 Equus cab
42	274.8	75.7	469	10	MATGFB1	X60296 M.auratus m
43	271.6	74.8	1597	10	AF191297	AF191297 Cavia por
44	271.2	74.7	374	4	AF349538	AF349538 Canis fam
45	266.8	73.5	336	6	A23758	A23758 TGF-beta3(4

ALIGNMENTS

RESULT 1

AX338213

LOCUS AX338213 1326 bp DNA linear PAT 09-JAN-2002

DEFINITION Sequence 1 from Patent WO0181404.

ACCESSION AX338213

VERSION AX338213.1 GI:18128750

KEYWORDS .

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1

AUTHORS Strober,W., Nakamura,K., Kitani,A. and Fuss,I.J.

TITLE Inducible plasmid vector encoding tgf-_g(b) and uses thereof

JOURNAL Patent: WO 0181404-A 1 01-NOV-2001;

THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)

FEATURES

Location/Qualifiers

source

1. .1326

/organism="Sus scrofa"

/mol_type="genomic DNA"

/db_xref="taxon:9823"

CDS

16. .1188

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD20538.1"

/db_xref="GI:18128751"

/translation="MAPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK
RIEAIRGQILSKLRLASPPSQGDVPPGGLPEAVLALYNSTRDRVAGESVEPEPEPEAD
YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLRL
KLKVEQHVELYQKYSNDSWRYLSNRLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR
LSAHSSSDSKDNTLHVEINGFNSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"

BASE COUNT 263 a 438 c 392 g 233 t

ORIGIN

Query Match 84.5%; Score 306.8; DB 6; Length 1326;

Best Local Similarity 99.4%; Pred. No. 5.2e-52;

Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      879 CACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 938

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db      939 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 998

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAAACC 233
      |||||
Db      999 CATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAAACC 1058

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db     1059 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 1118

Qy      294 CTACGTGGGCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC GTTCCTGCAA 353
      |||||
Db     1119 CTACGTGGGCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC GTTCCTGCAA 1178

Qy      354 GTGCAGCTGA 363
      |||||
Db     1179 GTGCAGCTGA 1188

```

RESULT 2

SSTGFBR

LOCUS SSTGFBR 1605 bp mRNA linear MAM 27-MAR-1995

DEFINITION Porcine mRNA for transforming growth factor-beta (TFG) precursor.

ACCESSION Y00111

VERSION Y00111.1 GI:2129

KEYWORDS transforming growth factor-beta.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 1605)

AUTHORS Derynck,R. and Rhee,L.

TITLE Sequence of the porcine transforming growth factor-beta precursor

JOURNAL Nucleic Acids Res. 15 (7), 3187 (1987)

MEDLINE 87174844

PUBMED 3470708

FEATURES

```

source          Location/Qualifiers
                1. .1605
                /organism="Sus scrofa"
                /mol_type="mRNA"
                /db_xref="taxon:9823"
                /tissue_type="ovary"
                /clone_lib="lambda gt10"
CDS             404. .1576
                /note="TFG-beta precursor (AA 1-390)"
                /codon_start=1
                /protein_id="CAA68291.1"
                /db_xref="GI:2130"
                /db_xref="SWISS-PROT:P07200"
                /translation="MPPSGLRLLPLLLPLLWLLVLT PGRPAAGLSTCKTIDMELVKRK
                RIEAIRGQILSKRLASPPSQGDVPPGP LPEAVLALYNSTRDRVAGESVEPEPEPEAD

```

```

YYAKEVTRVLMLESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAI EGFR
LSAHCSCDSKDNTLHVEINGFNSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
mat_peptide      1238.  .1573
                  /product="mature TFG-beta"
misc_feature      647.  .655
                  /note="pot. N-glycosylation site"
misc_feature      809.  .817
                  /note="pot. N-glycosylation site"
BASE COUNT      298 a      572 c      437 g      298 t
ORIGIN

```

```

Query Match      84.5%;   Score 306.8;   DB 4;   Length 1605;
Best Local Similarity 99.4%;   Pred. No. 5.1e-52;
Matches 308;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      1267 CACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 1326

Qy      114 GAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db      1327 GAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 1386

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAAACCC 233
      |||||
Db      1387 CATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAAACCC 1446

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db      1447 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 1506

Qy      294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTTCCTGCAA 353
      |||||
Db      1507 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTTCCTGCAA 1566

Qy      354 GTGCAGCTGA 363
      |||||
Db      1567 GTGCAGCTGA 1576

```

RESULT 3

GGTGFB1

```

LOCUS      GGTGFB1      1750 bp      mRNA      linear      MAM 27-MAR-1996
DEFINITION Porcine mRNA for transforming growth factor-beta 1.
ACCESSION  X12373
VERSION    X12373.1  GI:63808
KEYWORDS   transforming growth factor-beta 1.
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 1750)
AUTHORS    Jakowlew,S.B., Dillard,P.J., Sporn,M.B. and Roberts,A.B.
TITLE      Nucleotide sequence of chicken transforming growth factor-beta 1

```

(TGF-beta 1)

JOURNAL Nucleic Acids Res. 16 (17), 8730 (1988)

MEDLINE 88335639

PUBMED 3166520

REFERENCE 2 (bases 1 to 1750)

AUTHORS Jakowlew,S.B.

TITLE Direct Submission

JOURNAL Submitted (14-JUL-1988) Jakowlew S.B., National Institute of health, National Cancer Institute, Laboratory of Chemoprevention, Building 41, Room B902, Bethesda, Maryland 20892, USA

COMMENT The submitters believe that the chicken cDNA library was contaminated with porcine cDNA, and that the sequence is infact porcine TGF-beta-1. 27-MAR-1996.

FEATURES Location/Qualifiers

source 1..1750

/organism="Sus scrofa"

/mol_type="mRNA"

/strain="white leghorn"

/db_xref="taxon:9823"

/clone="pTGFB-ChX119"

/cell_type="chondrocyte"

gene 1..1750

/gene="TGF-beta 1"

5'UTR 1..446

/gene="TGF-beta 1"

CDS 447..1622

/gene="TGF-beta 1"

/codon_start=1

/product="transforming growth factor"

/protein_id="CAA30933.1"

/db_xref="GI:63809"

/db_xref="SWISS-PROT:P07200"

/translation="MPPSGPGLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK
 RIEAIRGQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD
 YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL
 KLKVEQHVELYQKYSNDSWGYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR
 LSAHCSCDSKDNTHLVEINAGFNSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSR
 HRRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD
 TQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYVYVGRKPKVEQLSNMIVRSCKCS"

mat_peptide 1285..1619

/gene="TGF-beta 1"

3'UTR 1620..1750

/gene="TGF-beta 1"

BASE COUNT 325 a 627 c 479 g 319 t

ORIGIN

Query Match 84.5%; Score 306.8; DB 4; Length 1750;

Best Local Similarity 99.4%; Pred. No. 5.1e-52;

Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113

Db 1313 CACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 1372

Qy 114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173

Db 1373 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 1432

```

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAAACC 233
        |||
Db      1433 CATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAACC 1492

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
        |||
Db      1493 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 1552

Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
        |||
Db      1553 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 1612

Qy      354 GTGCAGCTGA 363
        |||
Db      1613 GTGCAGCTGA 1622

```

RESULT 4

AF461808

LOCUS AF461808 2221 bp mRNA linear MAM 03-JAN-2002

DEFINITION Sus scrofa transforming growth factor beta 1 (TGFB1) mRNA, complete cds.

ACCESSION AF461808

VERSION AF461808.1 GI:18042250

KEYWORDS .

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Polymorphism in the porcine transforming growth factor beta 1 gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Direct Submission

JOURNAL Submitted (20-DEC-2001) Institute of Animal Breeding Science, University of Bonn, Endenicher Allee 15, Bonn 53115, Germany

FEATURES Location/Qualifiers

source

1. .2221
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/chromosome="6"

gene

1. .2221
/gene="TGFB1"

CDS

1. .1173
/gene="TGFB1"
/note="cytokine"
/codon_start=1
/product="transforming growth factor beta 1"
/protein_id="AAL57902.1"
/db_xref="GI:18042251"
/translation="MPPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK
RIEAI RQILSKLR L A S P P S Q G D V P P G P L P E A V L A L Y N S T R D R V A G E S V E P E P E P E A D
Y Y A K E V T R V L M V E S G N Q I Y D K F K G T P H S L Y M L F N T S E L R E A V P E P V L L S R A E L R L L R L

```

KLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTVGVVRQWLTRREAIEGFR
LSAHCSCDSKDNTHLVEINGFNSGRRGDLATIHGMNRPFLLLMTPLERAQHLHSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"
3'UTR 1174. .2221
/gene="TGFB1"
BASE COUNT 504 a 616 c 649 g 452 t
ORIGIN

```

```

Query Match      84.5%;  Score 306.8;  DB 4;  Length 2221;
Best Local Similarity 99.4%;  Pred. No. 4.9e-52;
Matches 308;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      |||
Db      864 CACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 923
      |||

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||
Db      924 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 983
      |||

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAAACC 233
      |||
Db      984 CATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAAACC 1043
      |||

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||
Db      1044 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 1103
      |||

Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
      |||
Db      1104 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 1163
      |||

Qy      354 GTGCAGCTGA 363
      |||
Db      1164 GTGCAGCTGA 1173

```

RESULT 5

PIGTGFB1A

```

LOCUS      PIGTGFB1A 3206 bp mRNA linear MAM 31-MAR-1995
DEFINITION Sus scrofa transforming growth factor beta-1 mRNA, complete cds.
ACCESSION  M23703
VERSION    M23703.1 GI:755044
KEYWORDS   transforming growth factor-beta-1.
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 3206)
AUTHORS    Kondaiah,P., Van Obberghen-Schilling,E., Ludwig,R.L., Dhar,R.,
            Sporn,M.B. and Roberts,A.B.
TITLE      cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.
            Evidence for alternate splicing and polyadenylation
JOURNAL    J. Biol. Chem. 263 (34), 18313-18317 (1988)
MEDLINE    89054010
PUBMED     2461367

```

COMMENT On Apr 1, 1995 this sequence version replaced gi:341017.
Original source text: Sus scrofa (strain miniature swine) cDNA to
mRNA.

Query Match 84.5%; Score 306.8; DB 4; Length 3206;
Best Local Similarity 99.4%; Pred. No. 4.7e-52;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db |||||
2069 GTGCAGCTGA 2078

RESULT 6

AR036686

LOCUS AR036686 339 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 29 from patent US 5800811.

ACCESSION AR036686

VERSION AR036686.1 GI:5954542

KEYWORDS .

SOURCE Unknown.

 ORGANISM Unknown.

 Unclassified.

REFERENCE 1 (bases 1 to 339)

 AUTHORS Hall,F.L., Nimni,M.E., Tuan,T.-L., Wu,L. and Cheung,D.T.

 TITLE Artificial skin prepared from coclagen matrix containing
 transforming growth factor-.beta. having a collagen binding site

 JOURNAL Patent: US 5800811-A 29 01-SEP-1998;

FEATURES Location/Qualifiers

 source 1. .339

 /organism="unknown"

BASE COUNT 66 a 113 c 100 g 60 t

ORIGIN

Query Match 77.5%; Score 281.2; DB 6; Length 339;

Best Local Similarity 94.2%; Pred. No. 8.9e-47;

Matches 292; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113

 || |||||

Db 30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy 114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173

 |||||

Db 90 GAAGTGGATCCATGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTA 149

Qy 174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAAACC 233

 |||||

Db 150 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAAACC 209

Qy 234 GGGCGCGTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293

 |||||

Db 210 GGGCGCCTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 269

Qy 294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353

 |||||

Db 270 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 329

Qy 354 GTGCAGCTGA 363

 |||||

Db 330 GTGCAGCTGA 339

RESULT 7

AX455100

LOCUS AX455100 489 bp DNA linear PAT 06-JUL-2002

DEFINITION Sequence 167 from Patent WO0208453.
 ACCESSION AX455100
 VERSION AX455100.1 GI:21714285
 KEYWORDS .
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 REFERENCE 1
 AUTHORS Farr,S.B., Pickett,G.G., Neft,R.E. and Dunn,R.T.
 TITLE Canine toxicity genes
 JOURNAL Patent: WO 0208453-A 167 31-JAN-2002;
 Phase-1 Molecular Toxicology (US)

FEATURES Location/Qualifiers
 source 1..489
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /db_xref="taxon:9615"
 BASE COUNT 84 a 183 c 145 g 76 t 1 others
 ORIGIN

Query Match 77.5%; Score 281.2; DB 6; Length 489;
 Best Local Similarity 94.2%; Pred. No. 8.5e-47;
 Matches 292; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      |||||||
Db     101 CACGGAGAAGAACTGCTGCGTCCGGCAGCTCTACATTGACTTCCGCAAGGATCTGGGCTG 160

Qy     114 GAAGTGGATTGATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||||
Db     161 GAAGTGGATCCATGAGCCCAAGGGTTACCAAGCTAACTTCTGCCTGGGGCCCTGCCCCCTA 220

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAACCC 233
      |||
Db     221 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCACAACCC 280

Qy     234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||||
Db     281 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 340

Qy     294 CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
      |||||||
Db     341 CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTGGAACATGATCGTGCGTTCCTGCAA 400

Qy     354 GTGCAGCTGA 363
      |||
Db     401 GTGCAGCTGA 410
  
```

RESULT 8
 OATGFB1
 LOCUS OATGFB1 1173 bp mRNA linear MAM 18-APR-1995
 DEFINITION O.aries mRNA for transforming growth factor-beta I.
 ACCESSION X76916
 VERSION X76916.1 GI:496648
 KEYWORDS TGF-beta 1; transforming growth factor-beta 1.

```

SOURCE          Ovis aries (sheep)
ORGANISM        Ovis aries
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
                Bovidae; Caprinae; Ovis.
REFERENCE       1
AUTHORS         Woodall,C.J., McLaren,L.J. and Watt,N.J.
TITLE           Sequence and chromosomal localisation of the gene encoding ovine
                latent transforming growth factor-beta 1
JOURNAL         Gene 150 (2), 371-373 (1994)
MEDLINE         95121932
PUBMED          7821809
REFERENCE       2 (bases 1 to 1173)
AUTHORS         Woodall,C.
TITLE           Direct Submission
JOURNAL         Submitted (24-DEC-1993) C. Woodall, Univ. of Edinburgh, Dept. of
                Veterinary Pathology, Sc. of Vet. Studies, Univ. of Edinburgh,
                Edinburgh EH9 1QH, UK
FEATURES        Location/Qualifiers
                source                1. .1173
                                     /organism="Ovis aries"
                                     /mol_type="mRNA"
                                     /db_xref="taxon:9940"
                CDS                   1. .1173
                                     /codon_start=1
                                     /product="transforming growth factor-beta 1"
                                     /protein_id="CAA54242.1"
                                     /db_xref="GI:496649"
                                     /db_xref="SWISS-PROT:P50414"
                                     /translation="MPPSGLRLLPLLLPLLWLLMLTPGRPVPAGLSTCKTIDMELVKRK
                mat_peptide            835. .1170
                                     /product="transforming growth factor-beta 1"
BASE COUNT      245 a      378 c      336 g      214 t
ORIGIN

```

```

Qy      234 GGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1044 GGGCGCATCGGCGGCGCCGTGCTGCGTGCCCTCAGGCGCTGGAACCCCTGCCCATCGTGTA 1103

Qy      294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1104 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGTTGTCCAACATGATCGTGCGCTCCTGCAA 1163

Qy      354 GTGCAGCTGA 363
          ||||| |||||
Db      1164 GTGCAGCTGA 1173

```

RESULT 9

DOGTGFB1A

LOCUS DOGTGFB1A 1369 bp mRNA linear MAM 30-OCT-1994

DEFINITION Canine transforming growth factor-beta 1 (TGFB1) mRNA, complete cds.

ACCESSION L34956

VERSION L34956.1 GI:516071

KEYWORDS homologue; transforming growth factor-beta 1.

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 1369)

AUTHORS Manning,A.M., Auchampach,J.A., Drong,R.F. and Slightom,J.L.

TITLE Cloning of a canine cDNA homologous to human transforming growth factor-beta 1 (TGFbeta1)

JOURNAL Unpublished (1994)

COMMENT Original source text: Canis familiaris adult jugular vein endothelial cDNA to mRNA.

FEATURES Location/Qualifiers

```

source      1..1369
             /organism="Canis familiaris"
             /mol_type="mRNA"
             /db_xref="taxon:9615"
             /cell_type="LPS-activated"
             /tissue_type="jugular vein endothelial"
             /dev_stage="adult"

```

```

gene        1..1369
             /gene="TGFB1"

```

```

5'UTR       1..57
             /gene="TGFB1"

```

```

CDS         58..1230
             /gene="TGFB1"
             /function="anti-inflammatory agent"
             /note="precursor"
             /codon_start=1
             /product="transforming growth factor-beta 1"
             /protein_id="AAA51458.1"
             /db_xref="GI:516072"
             /translation="MPPSGLRLLPLLLPLLRLLVLTTPGRPAAGLSTCKTIDMELVKRK
RIEAIRGQILSKLRLLSSPPSQGEVPPVPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
YYAKEVTRVLMVENTNKIYEKVKKSPHSIYMLFNTSELREAVPEPVLLSRAELRLLRL
KLKAEQHVELYQKYSNDSWRYLNRLLAPSDTPEWLSFDVTGVVRQWLSHGGEVEGFR
LSAHCSCDSKDNTLQVDINGFSSSRRGDLATIHGMRPFLLLLMATPLERAQHLHSSRQ

```

```

RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"
mat_peptide      1039.  .1227
                  /gene="TGFB1"
                  /product="transforming growth factor-beta 1"
3'UTR            1231.  .1369
                  /gene="TGFB1"
BASE COUNT      264 a      473 c      415 g      216 t      1 others
ORIGIN

```

```

Query Match      77.5%;  Score 281.2;  DB 4;  Length 1369;
Best Local Similarity 94.2%;  Pred. No. 7.5e-47;
Matches 292;  Conservative 0;  Mismatches 18;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      921 CACGGAGAAGAACTGCTGCGTCCGGCAGCTCTACATTGACTTCCGCAAGGATCTGGGCTG 980

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      981 GAAGTGGATCCATGAGCCCAAGGGTTACACGCTAACTTCTGCCTGGGGCCCTGCCCCCTA 1040

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTCTGGCTCTGTACAACCAGCACAAACC 233
      ||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db      1041 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTCTGGCCCTGTACAACCAGCACAAACC 1100

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1101 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 1160

Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTTCCTGCAA 353
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1161 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTTCCTGCAA 1220

Qy      354 GTGCAGCTGA 363
      |||||||||
Db      1221 GTGCAGCTGA 1230

```

RESULT 10

AGMTGFB

LOCUS AGMTGFB 1561 bp mRNA linear PRI 27-APR-1993

DEFINITION Simian transforming growth factor-beta (TGF) mRNA, complete cds.

ACCESSION M16658

VERSION M16658.1 GI:176552

KEYWORDS growth factor; transforming growth factor-beta.

SOURCE Cercopithecus aethiops (African green monkey)

ORGANISM Cercopithecus aethiops

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Cercopithecus.

REFERENCE 1 (bases 1 to 1561)

AUTHORS Sharples,K., Plowman,G.D., Rose,T.M., Twardzik,D.R. and
Purchio,A.F.

TITLE Cloning and sequence analysis of simian transforming growth
factor-beta cDNA

JOURNAL DNA 6 (3), 239-244 (1987)

RESULT 11

A18277

LOCUS A18277 339 bp mRNA linear PAT 17-MAY-1994

DEFINITION H.sapiens TGF-beta 1 gene seq ID No:1.

ACCESSION A18277

VERSION A18277.1 GI:513237

KEYWORDS .

SOURCE synthetic construct

ORGANISM synthetic construct

artificial sequences.

REFERENCE 1 (bases 1 to 339)

AUTHORS Cerletti,N., McMaster,G.K., Cox,D., Schmitz,A. and Meyhack,B.

TITLE Process for the production of biologically active protein (e.g. TGF)

JOURNAL Patent: EP 0433225-A 1 19-JUN-1991;
CIBA-GEIGY AG

FEATURES Location/Qualifiers

source

1. .339

/organism="synthetic construct"

/mol_type="mRNA"

/db_xref="taxon:32630"

CDS

1. .339

/note="Protein sequence is in conflict with the conceptual translation"

/codon_start=1

/transl_table=11

/product="TGF-beta 1"

/protein_id="CAA01385.1"

/db_xref="GI:4529903"

/translation="ALDTNYCFSSSTEKNCCVRLYIDFRKDLGWKWIHEPKGYHANFC
LGPCPYIWSLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYVGRKPKVEQLS
NMIVRSCKCS"

BASE COUNT 66 a 114 c 100 g 59 t

ORIGIN

Query Match 77.0%; Score 279.6; DB 6; Length 339;

Best Local Similarity 93.9%; Pred. No. 1.9e-46;

Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113

|| |||||

Db 30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy 114 GAAGTGGATTTCATGAACCCAAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173

||||| || || |||||

Db 90 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTA 149

Qy 174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAACCC 233

||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 150 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCC 209

Qy 234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 210 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 269

Qy 294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 353

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 270 CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTCCTGCAA 329

Qy 354 GTGCAGCTGA 363
 |||||

Db 330 GTGCAGCTGA 339

RESULT 12

A23751

LOCUS A23751 339 bp mRNA linear PAT 25-JAN-1995

DEFINITION TGF-beta1 coding region.

ACCESSION A23751

VERSION A23751.1 GI:825585

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 339)

AUTHORS McMaster,G.K., Cox,D., Cerletti,N. and Kuhla,J.

TITLE Novel hybrid transforming growth factors

JOURNAL Patent: EP 0542679-A 1 19-MAY-1993;

CIBA-GEIGY AG

FEATURES Location/Qualifiers

source 1..339
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

BASE COUNT 66 a 114 c 100 g 59 t

ORIGIN

Query Match 77.0%; Score 279.6; DB 6; Length 339;
 Best Local Similarity 93.9%; Pred. No. 1.9e-46;
 Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGGCTG 113
 ||||||||||||||||||||||||||||||||||||||||||||

Db 30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGGCTG 89

Qy 114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
 ||||||||| || || ||||||||||||||||||||||||||||

Db 90 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCCCTA 149

Qy 174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTTGGCTCTGTACAACCAGCACAACCC 233
 ||| |||||||| |||| ||||||||||||||||||||||||

Db 150 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGGCCCTGTACAACCAGCATAACCC 209

Qy 234 GGGCGCGTTCGGCGGCGCGGTGCTGCGTGCCGCGGCGCTGGAGCCACTGCCCATCGTGTA 293
 ||||| ||||||||||||||||||||||||||||||||||||

Db 210 GGGCGCCTTCGGCGGCGCGGTGCTGCGTGCCGCGGCGCTGGAGCCGCTGCCCATCGTGTA 269

Qy 294 CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTCCTGCAA 353
 ||||||||||||||||||||||||||||||||||||||||

Db 270 CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTCCTGCAA 329

Qy 354 GTGCAGCTGA 363
 |||||||

Db 330 GTGCAGCTGA 339

RESULT 13

A48549

LOCUS A48549 339 bp DNA linear PAT 07-MAR-1997

DEFINITION Sequence 1 from Patent WO9603432.

ACCESSION A48549

VERSION A48549.1 GI:2302319

KEYWORDS .

SOURCE unidentified

ORGANISM unidentified

unclassified.

REFERENCE 1 (bases 1 to 339)

AUTHORS Cerletti,N.

TITLE NOVEL PROCESS FOR THE PRODUCTION OF BIOLOGICALLY ACTIVE DIMERIC
PROTEI

JOURNAL Patent: WO 9603432-A 1 08-FEB-1996;

CIBA GEIGY AG (CH)

COMMENT Other publication AU 3109595 960222.

FEATURES Location/Qualifiers

source 1. .339

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

CDS 1. .336

/note="Protein sequence is in conflict with the conceptual
translation"

/codon_start=1

/product="HUMAN TGF-BETA1"

/protein_id="CAA03113.1"

/db_xref="GI:2302320"

/translation="ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFC
LGPCPYIWSLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYVGRKPKVEQLS
NMIVRSCKCS"

BASE COUNT 66 a 114 c 100 g 59 t

ORIGIN

Query Match 77.0%; Score 279.6; DB 6; Length 339;

Best Local Similarity 93.9%; Pred. No. 1.9e-46;

Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113

|||||

Db 30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy 114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173

|||||

Db 90 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCCTA 149

Qy 174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAACCC 233

|||||

Db 150 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCC 209

Qy 234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293

|||||

Db 210 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 269


```

Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
          |||
Db      270 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAA 329

Qy      354 GTGCAGCTGA 363
          |||
Db      330 GTGCAGCTGA 339

```

RESULT 14

A48563

LOCUS A48563 339 bp DNA linear PAT 07-MAR-1997

DEFINITION Sequence 1 from Patent WO9603433.

ACCESSION A48563

VERSION A48563.1 GI:2302333

KEYWORDS .

SOURCE unidentified

ORGANISM unidentified

unclassified.

REFERENCE 1 (bases 1 to 339)

AUTHORS Cerletti,N.

TITLE NEW PROCESS FOR THE PRODUCTION OF BIOLOGICALLY ACTIVE PROTEIN

JOURNAL Patent: WO 9603433-A 1 08-FEB-1996;

CIBA GEIGY AG (CH)

COMMENT Other publication AU 3109695 960222.

FEATURES

source

Location/Qualifiers

1. .339

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

CDS

1. .336

/note="Protein sequence is in conflict with the conceptual translation"

/codon_start=1

/product="HUMAN TGF-BETA1"

/protein_id="CAA03120.1"

/db_xref="GI:2302334"

/translation="ALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFC LGPCPYIWSLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLS

NMIVRSCKCS"

BASE COUNT 66 a 114 c 100 g 59 t

ORIGIN

Query Match 77.0%; Score 279.6; DB 6; Length 339;

Best Local Similarity 93.9%; Pred. No. 1.9e-46;

Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
          ||
Db      30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy     114 GAAGTGGATTTCATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
          |||
Db      90 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCTA 149

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAAACC 233

```

Db	150	CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCC	209
Qy	234	GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA	293
Db	210	GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA	269
Qy	294	CTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA	353
Db	270	CTACGTGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAA	329
Qy	354	GTGCAGCTGA	363
Db	330	GTGCAGCTGA	339

RESULT 15

```

I56854
LOCUS           I56854                   339 bp      DNA          linear      PAT 07-OCT-1997
DEFINITION     Sequence 1 from patent US 5650494.
ACCESSION      I56854
VERSION        I56854.1  GI:2477267
KEYWORDS       .
SOURCE         Unknown.
               ORGANISM   Unknown.
               Unclassified.
REFERENCE      1  (bases 1 to 339)
               AUTHORS    Cerletti,N., McMaster,G.Kent., Cox,D., Schmitz,A. and Meyhack,B.
               TITLE      Process for refolding recombinantly produced TGF-.beta.-like
                           proteins
               JOURNAL     Patent: US 5650494-A 1 22-JUL-1997;
FEATURES       Location/Qualifiers
               source      1. .339
                           /organism="unknown"
BASE COUNT     66 a      114 c      100 g      59 t
ORIGIN

```

Query Match 77.0%; Score 279.6; DB 6; Length 339;
Best Local Similarity 93.9%; Pred. No. 1.9e-46;
Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

[illegible]

```
Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
```

17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	306.8	84.5	1326	24	AAD22696	Porcine transformi	
2	281.2	77.5	339	18	AAT42771	TGF-betal active f	
3	281.2	77.5	339	20	AAV99375	cDNA encoding a tr	
4	281.2	77.5	489	24	ABL99528	Target canine gene	
5	281.2	77.5	1303	11	AAQ09317	Monkey transformin	
6	281.2	77.5	1559	13	AAQ20289	Sequence encoding	
7	281.2	77.5	1561	11	AAQ03268	Simian transformin	
8	281.2	77.5	1571	11	AAQ03269	Human transforming	
9	279.6	77.0	339	17	AAT17235	Human TGF-beta 1 c	
10	279.6	77.0	339	20	AAX15245	cDNA encoding the	
11	279.6	77.0	650	24	ABK84023	Human cDNA differe	
12	279.6	77.0	650	24	ABL68818	Kidney cancer rela	
13	279.6	77.0	1176	24	ABZ35738	Human TGF beta 1 p	
14	279.6	77.0	1176	24	ABX09981	Human TGFbetal DNA	
15	279.6	77.0	1176	24	ABV78162	Human TGF beta 1 D	
16	279.6	77.0	1176	24	ABL91703	Human polynucleoti	
17	279.6	77.0	1176	25	ABV75391	TGFB1 Arg25Pro pol	
18	279.6	77.0	1176	25	ABV75392	TGFB1 Arg25Pro pol	
19	279.6	77.0	1560	9	AAN81084	Coding sequence of	
20	279.6	77.0	1560	11	AAQ03508	Simian Transformin	
21	279.6	77.0	1569	9	AAN81085	Coding sequence of	
22	279.6	77.0	1821	12	AAQ13392	Human pro-TGF-beta	
23	279.6	77.0	2537	7	AAN60972	Sequence encoding	
24	279.6	77.0	2537	11	AAQ03301	cDNA encoding huma	
25	279.6	77.0	2537	11	AAQ02814	Sequence of pre-TG	
26	279.6	77.0	2537	15	AAQ56923	Human pre-TGF-beta	
27	279.6	77.0	2537	17	AAT15720	Pre-transforming g	
28	279.6	77.0	2537	19	AAV52933	Human pre-transfor	
29	279.6	77.0	2742	22	AAI58342	Human polynucleoti	
30	279.6	77.0	2745	16	AAT05876	cDNA encoding tran	
31	279.6	77.0	2745	22	AAH28216	Nucleotide sequenc	
32	279.6	77.0	3541	17	AAT16516	Collagen A1/TGF-be	
33	279.6	77.0	3541	21	AAA12498	cDNA encoding a ch	
34	279.6	77.0	4105	15	AAQ55624	TGFbetal 5'-UTR-CD	
35	278	76.6	339	12	AAQ11993	Transforming Growt	
36	278	76.6	339	17	AAT15462	Human transforming	
37	278	76.6	1266	24	ABK12871	Chimeric TGF-B-SA	
38	278	76.6	1569	11	AAQ03509	Human Transforming	

39	278	76.6	2527	25	ABQ76674	Androgen receptor
40	275.4	75.9	345	14	AAQ41599	Mature human Trans
41	266.8	73.5	336	14	AAQ41606	Transforming Growt
42	262.2	72.2	1561	11	AAQ04908	Sequence encoding
43	262.2	72.2	1561	13	AAQ29177	TGF-beta 1/beta 2
44	253.2	69.8	875	23	AAS70979	DNA encoding novel
45	248	68.3	336	14	AAQ41604	Transforming Growt

ALIGNMENTS

RESULT 1

AAD22696

ID AAD22696 standard; cDNA; 1326 BP.

XX

AC AAD22696;

XX

DT 26-FEB-2002 (first entry)

XX

DE Porcine transforming growth factor beta 1 (TGF-beta1) cDNA.

XX

KW Porcine; transforming growth factor beta 1; TGF-beta1; gene therapy;
 KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;
 KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; ss.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT CDS 16..1188

FT /*tag= a

FT /product= "Porcine TGF-beta1 mutant protein"

XX

PN WO200181404-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12980.

XX

PR 20-APR-2000; 2000US-199014P.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Strober W, Nakamura K, Kitani A, Fuss IJ;

XX

DR WPI; 2002-026155/03.

DR P-PSDB; AAE13596.

XX

PT Composition for treating autoimmune diseases e.g. inflammatory bowel
 PT disease in humans, comprises vector containing transforming growth
 PT factor-beta under the control of inducible promoter -

XX

PS Claim 1; Fig 1; 78pp; English.

XX

CC The invention relates to a composition containing a vector comprising a
 CC gene encoding a regulatory transcription factor under the control of a

XX
 KW Transforming growth factor-beta fusion protein; wound healing;
 KW artificial skin; surgery recovery time; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..339
 FT /*tag= a
 FT /function= TGF active fragment
 XX
 PN WO9639430-A1.
 XX
 PD 12-DEC-1996.
 XX
 PF 05-JUN-1996; 96WO-US08973.
 XX
 PR 06-JUN-1995; 95US-0470837.
 XX
 PA (CHEU/) CHEUNG D T.
 PA (HALL/) HALL F L.
 PA (NIMN/) NIMNI M E.
 PA (TUAN/) TUAN T.
 PA (WULL/) WU L.
 XX
 PI Cheung DT, Hall FL, Nimni ME, Tuan T, Wu L;
 XX
 DR WPI; 1997-043065/04.
 DR P-PSDB; AAW08173.
 XX
 PT Prepn. of transforming growth factor-beta fusion protein - useful to
 PT reduce surgery recovery time and to prepare artificial skin
 XX
 PS Disclosure; Page 44-45; 59pp; English.
 XX
 CC A novel transforming growth factor-beta (TGF-beta) fusion protein
 CC comprises a purification tag and a TGF active fragment. The present
 CC sequence encodes a specifically claimed TGF active fragment, TGF-beta1.
 CC Additionally, the fusion protein may comprise proteinase-sensitive
 CC linker sites and binding domain so the protein sequence may contain
 CC some or all of the following elements: purification tag:proteinase
 CC site:ECM binding site:proteinase site:TGF-beta. TGF-beta promotes
 CC wound healing, and the fusion protein can be used to reduce surgery
 CC recovery time and in the preparation of artificial skin. The inclusion
 CC of a purification tag facilitates purification of the fusion protein.
 CC The proteinase site is included to permit cleavage and release of the
 CC purification tag after purification if desired. The extracellular
 CC matrix binding site facilitates delivery of the fusion protein to the
 CC desired site of action. Delivery of the TGF-beta to the site to be
 CC treated reduces the amount of TGF-beta required to be administered to
 CC be effective and reduces the concentration of circulating TGF-beta
 CC which may result in undesirable effects.
 XX
 SQ Sequence 339 BP; 66 A; 113 C; 100 G; 60 T; 0 other;

Query Match 77.5%; Score 281.2; DB 18; Length 339;
 Best Local Similarity 94.2%; Pred. No. 3e-58;

Matches 292; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy     114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db      90 GAAGTGGATCCATGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTA 149

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAACCC 233
      |||||
Db     150 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCC 209

Qy     234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db     210 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 269

Qy     294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 353
      |||||
Db     270 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 329

Qy     354 GTGCAGCTGA 363
      |||||
Db     330 GTGCAGCTGA 339

```

RESULT 3

AAV99375

ID AAV99375 standard; cDNA; 339 BP.

XX

AC AAV99375;

XX

DT 25-MAR-1999 (first entry)

XX

DE cDNA encoding a transforming growth factor beta active fragment.

XX

KW Proteinase site; bone morphogenetic fusion protein; bone binding site;

KW bone morphogenetic protein; transforming growth factor beta;

KW active fragment; wound healing; bone growth; purification tag; ds.

XX

OS Homo sapiens.

XX

PN WO9855137-A1.

XX

PD 10-DEC-1998.

XX

PF 02-JUN-1998; 98WO-US11189.

XX

PR 03-JUN-1997; 97US-0868452.

XX

PA (HALL/) HALL F L.

PA (HANB/) HAN B.

PA (NIMN/) NIMNI M E.

PA (SHOR/) SHORS E C.

PA (WULL/) WU L.

XX

AC ABL99528;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Target canine gene TGFB1.
 XX
 KW Canine gene array; toxicological response; ss.
 XX
 OS Canis sp.
 XX
 PN WO200208453-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 23-JUL-2001; 2001WO-US23311.
 XX
 PR 21-JUL-2000; 2000US-220057P.
 XX
 PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY.
 XX
 PI Farr SB, Pickett GG, Neft RE, Dunn RT;
 XX
 DR WPI; 2002-217063/27.
 XX
 PT Identifying toxicologically relevant canine gene to determine
 PT toxicological responses of agents, by obtaining and comparing gene
 PT expression profiles of untreated canine cells and canine cells treated
 PT with an agent -
 XX
 PS Disclosure; Page 67; 140pp; English.
 XX
 CC This invention relates to identifying a toxicologically relevant canine
 CC gene and the generation of an array of toxicologically relevant
 CC canine genes. The gene array is useful for obtaining a gene expression
 CC profile, by exposing a population of cells to an agent, obtaining cDNA
 CC from the population of cells, labeling the cDNA, and contacting the cDNA
 CC with the gene array. The relevant gene is useful for making and using
 CC arrays to determine toxicological responses to various agents, and also
 CC useful for identifying novel gene sequences and novel canine genes.
 CC The method for analysing toxicological responses using the canine
 CC gene array is rapid and efficient. The present sequence is related
 CC to the canine gene array.
 XX
 SQ Sequence 489 BP; 84 A; 183 C; 145 G; 76 T; 1 other;

 Query Match 77.5%; Score 281.2; DB 24; Length 489;
 Best Local Similarity 94.2%; Pred. No. 3.2e-58;
 Matches 292; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

 Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
 || |||||
 Db 101 CACGGAGAAGAACTGCTGCGTCCGGCAGCTCTACATTGACTTCCGCAAGGATCTGGGCTG 160

 Qy 114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
 |||||
 Db 161 GAAGTGGATCCATGAGCCCAAGGGTTACACGCTAACTTCTGCCTGGGGCCCTGCCCTA 220

[illegible]

RESULT 5

AAO09317

ID AAQ09317 standard; cDNA; 1303 BP.

XX

AC AAQ09317;

XX

DT 25-MAR-2003 (updated)

DT 12-AUG-1990 (first entry)

XX

DE Monkey transforming growth factor-beta cDNA.

XX

KW Transforming growth factor-beta; simian; psoriasis;

KW TGF-beta.

XX

OS Monkey.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	sig peptide	22..63
----	-------------	--------

```
FT /*tag= a
```

FT	mat peptide	836..1170
----	-------------	-----------

```
FT                               /*tag= b
```

FT /product=monkey transforming growth factor-beta

XX

PN EP353772-A.

XX

PD 07-FEB-1990.

XX

PF 04-AUG-1989; 89EP-0114458.

XX

PR 05-AUG-1988; 88US-0229133.

XX

PA (ONCO) ONCOGEN LP.

XX

PI Twardzik DR, Purchio AF, Ranchalis JE, Stevens V;

XX

DR WPI; 1990-038499/06.

DR P-PSDB; AAR03743.

XX

PT Inhibition of proliferation of epidermal cells -

PT used to treat psoriasis by contacting cells with compositions

PT containing transforming growth factor-beta.
XX
PS Disclosure; fig 1; 20pp; English.
XX
CC TGF-beta may be used in the treatment of hyperplasia
CC associated with acanthosis-categorised skin diseases, and
CC in alleviating psoriatic symptoms associated with cytokine-
CC induced phenomena. See also AAQ03268 and AAR03750.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1303 BP; 263 A; 442 C; 378 G; 217 T; 3 other;

Query Match 77.5%; Score 281.2; DB 11; Length 1303;
Best Local Similarity 94.2%; Pred. No. 3.8e-58;
Matches 292; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      867 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGCTG 926

Qy     114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db     927 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 986

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTCTGGCTCTGTACAACCAGCACAACCC 233
      |||||
Db     987 CATTTGGAGCCTTGACACGCAGTACAGCAAGGTCTCTGGCCCTGTACAACCAGCATAACCC 1046

Qy     234 GGGCGCGTTCGGCGGCGCGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db    1047 GGGCGCCTCGGCGGCGCGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCATCGTGTA 1106

Qy     294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
      |||||
Db    1107 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 1166

Qy     354 GTGCAGCTGA 363
      |||||
Db    1167 ATGCAGCTGA 1176
```

RESULT 6

AAQ20289

ID AAQ20289 standard; cDNA; 1559 BP.

XX

AC AAQ20289;

XX

DT 25-MAR-2003 (updated)

DT 16-APR-1992 (first entry)

XX

DE Sequence encoding simian transforming growth factor (TGF) beta-1.

XX

KW Hypertension therapy; hypotensive agent; blood pressure modulator;
KW ss.

XX

OS Monkey.

XX

FH	Key	Location/Qualifiers
FT	CDS	262..282
FT		/*tag= a
FT	sig_peptide	283..324
FT		/*tag= b
FT	CDS	325..1098
FT		/*tag= c
FT	mat_peptide	1099..1436
FT		/*tag= d

XX

PN WO9119513-A.

XX

PD 26-DEC-1991.

XX

PF 20-JUN-1991; 91WO-US04449.

XX

PR 20-JUN-1990; 90US-0541221.

XX

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX

PI Oleson FB, Comereski CR;

XX

DR WPI; 1992-024199/03.

DR

P-PSDB; AAR20124.

XX

PT Use of transforming growth factor (TGF)-beta and their

PT antagonists - for modulating blood pressure, for treating

PT hypertension and hypotension

XX

PS Disclosure; Fig 1; 42pp; English.

XX

CC A new method for treating hypertension comprises administering a

CC transforming growth factor (TGF)-beta to an individual at a dose

CC effective for lowering blood pressure; the TGF-beta may be e.g.

CC mature TGF-beta, TGF-beta2, a mature TGF-beta1/beta2 hybrid, TGF-

CC beta1 precursor, a latent TGF-beta2 precursor, hybrid TGF-beta1/TGF-

CC beta2 precursor, a latent TGF-beta1 complex or a latent TGF-beta2

CC complex.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX

SQ Sequence 1559 BP; 300 A; 546 C; 446 G; 267 T; 0 other;

Query Match 77.5%; Score 281.2; DB 13; Length 1559;

Best Local Similarity 94.2%; Pred. No. 4e-58;

Matches 292; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113

|| |||||

Db 1124 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGCTG 1183

Qy 114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173

|||||

Db 1184 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTGGGGCCCTGTCCCTA 1243

Qy 174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAACCC 233

||| |||||

Db 1244 CATTGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCC 1303

```

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
        |||||
Db      1304 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 1363

Qy      294 CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 353
        |||||
Db      1364 CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 1423

Qy      354 GTGCAGCTGA 363
        |||||
Db      1424 ATGCAGCTGA 1433

```

RESULT 7

AAQ03268

ID AAQ03268 standard; DNA; 1561 BP.

XX

AC AAQ03268;

XX

DT 25-MAR-2003 (updated)

DT 12-AUG-1990 (first entry)

XX

DE Simian transforming growth factor-beta cDNA.

XX

KW Transforming growth factor-beta; psoriasis; TGF-beta; ss.

XX

OS Monkey.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	sig_peptide	283..324
----	-------------	----------

FT		/*tag= a
----	--	----------

FT	mat_peptide	1096..1431
----	-------------	------------

FT		/*tag= b
----	--	----------

FT		/product=human transforming growth factor-beta
----	--	--

XX

PN EP353772-A.

XX

PD 07-FEB-1990.

XX

PF 04-AUG-1989; 89EP-0114458.

XX

PR 05-AUG-1988; 88US-0229133.

XX

PA (ONCO) ONCOGEN LP.

XX

PI Twardzik DR, Purchio AF, Ranchalis JE, Stevens V;

XX

DR WPI; 1990-038499/06.

DR P-PSDB; AAR03743.

XX

PT Inhibition of proliferation of epidermal cells -

PT used to treat psoriasis by contacting cells with compositions

PT containing transforming growth factor-beta.

XX

PS Disclosure; fig 1; 20pp; English.

XX

CC TGF-beta may be used in the treatment of hyperplasia
CC associated with acanthosis-categorised skin diseases, and
CC in alleviating psoriatic symptoms associated with cytokine-
CC induced phenomena. See also AAQ03269 and AAR03750.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX

SQ Sequence 1561 BP; 301 A; 547 C; 446 G; 267 T; 0 other;

Query Match 77.5%; Score 281.2; DB 11; Length 1561;
Best Local Similarity 94.2%; Pred. No. 4e-58;
Matches 292; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      1125 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGCTG 1184

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db      1185 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 1244

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAAACC 233
      |||
Db      1245 CATTTGGAGCCTGGACACGAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAAACC 1304

Qy      234 GGGCGCGTCCGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db      1305 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 1364

Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
      |||||
Db      1365 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 1424

Qy      354 GTGCAGCTGA 363
      |||||
Db      1425 ATGCAGCTGA 1434
```

RESULT 8

AAQ03269

ID AAQ03269 standard; DNA; 1571 BP.

XX

AC AAQ03269;

XX

DT 25-MAR-2003 (updated)

DT 12-AUG-1990 (first entry)

XX

DE Human transforming growth factor-beta cDNA.

XX

KW Transforming growth factor-beta; psoriasis;

KW TGF-beta.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT sig_peptide 22..63

FT /*tag= a

FT mat_peptide 836..1170

```

FT          /*tag=  b
FT          /product=human transforming growth factor-beta
XX
PN  EP353772-A.
XX
PD  07-FEB-1990.
XX
PF  04-AUG-1989;      89EP-0114458.
XX
PR  05-AUG-1988;      88US-0229133.
XX
PA  (ONCO ) ONCOGEN LP.
XX
PI  Twardzik DR,  Purchio AF,  Ranchalis JE,  Stevens V;
XX
DR  WPI; 1990-038499/06.
DR  P-PSDB; AAR03750.
XX
PT  Inhibition of proliferation of epidermal cells -
PT  used to treat psoriasis by contacting cells with compositions
PT  containing transforming growth factor-beta.
XX
PS  Disclosure; fig 1; 20pp; English.
XX
CC  TGF-beta may be used in the treatment of hyperplasia
CC  associated with acanthosis-categorised skin diseases, and
CC  in alleviating psoriatic symptoms associated with cytokine-
CC  induced phenomena. See also AAQ03268 and AAR03743.
CC  (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ  Sequence 1571 BP; 299 A; 563 C; 443 G; 266 T; 0 other;

```


Db |||||
1435 ATGCAGCTGA 1444

RESULT 9

AAT17235

ID AAT17235 standard; cDNA to mRNA; 339 BP.

XX

AC AAT17235;

XX

DT 17-JUL-1996 (first entry)

XX

DE Human TGF-beta 1 cDNA.

XX

KW Transforming growth factor type beta; TGF-beta 1;
KW protein renaturation; protein folding; ds.

XX

OS Homo sapiens.

XX

PN W09603433-A1.

XX

PD 08-FEB-1996.

XX

PF 12-JUL-1995; 95WO-EP02719.

XX

PR 25-JUL-1994; 94EP-0810439.

XX

PA (CIBA) CIBA GEIGY AG.

XX

PI Cerletti N;

XX

DR WPI; 1996-117000/12.

DR P-PSDB; AAR92773.

XX

PT Prodn. of dimeric biologically active transforming growth factor -
PT by refolding denatured monomer in detergent-free folding buffer
PT contg. specific organic solvent to improve yield

XX

PS Example 1B; Page 20-30; 54pp; English.

XX

CC The coding sequence (AAT17235) of human transforming growth factor
CC TGF-beta 1 (AAR92773) was cloned into plasmid pGEM-5ZF(+) (Promega)
CC and the construct used to transform E. coli Y1090. Subcloning in
CC pPLMu yielded plasmid pPLMu.hTGF-beta 1. Non-soluble, monomeric
CC TGF-beta 1 was recovered from E. coli LC 137/pPLMu.hTGF-beta 1 (DSM
CC 5656) transformants. A biologically active, dimeric form of
CC TGF-beta 1 was obtd. by refolding this monomer in detergent-free
CC buffer contg. DMSO and/or DMF. Dimers of TGF-beta 3 (AAR92772) and
CC TGF-beta 2 (AAR92774), and hybrid dimers (see also AAR92775-77), were
CC also produced.

XX

SQ Sequence 339 BP; 66 A; 114 C; 100 G; 59 T; 0 other;

Query Match 77.0%; Score 279.6; DB 17; Length 339;

Best Local Similarity 93.9%; Pred. No. 7.3e-58;

Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

[illegible]

AAX15245

XX

XX

DT 28-APR-1999 (first entry)

DE cDNA encoding the mature form of transforming growth factor-beta-1.

KW

KW bone repair; tissue repair; bone marrow protective agent;

KW ulcer; bed sore; ds.

OS

PN

PD

PF

XX

PR

YV

FA
VV

AA
BT

PI

DR WPI; 1999-083520/08.
 DR P-PSDB; AAW97091.
 XX
 PT Producing biologically active dimeric Transforming Growth
 PT Factor-beta - by refolding new monomeric Transforming Growth
 PT Factor-beta, useful for treatment of wounds and cancer
 XX
 PS Example 1; Page 28; 32pp; English.
 XX
 CC The present sequence encodes the mature form of transforming growth
 CC factor-beta-1. Dimeric, biologically active TGF-beta-like protein
 CC can be produced by subjecting the denatured monomeric form to refolding
 CC conditions. The new monomeric S-sulphonated TGF-beta-like protein is
 CC useful for the production of the dimeric, biologically active
 CC TGF-beta-like protein, which is useful for the treatment of wounds
 CC (surface or internal) and cancer in a mammal, in bone and tissue
 CC repair, as a bone marrow protective agent, a mediator of
 CC cardioprotection, for the production of an anti-inflammatory or
 CC immunosuppressive preparation. Treatment is useful for animals,
 CC especially humans, and wound treatment (e.g. ulcers, bed sores etc.) is
 CC particularly useful for the elderly.
 CC (Updated on 20-MAR-2003 to correct PF field.)
 CC (Updated on 20-MAR-2003 to correct PR field.)
 XX
 SQ Sequence 339 BP; 66 A; 114 C; 100 G; 59 T; 0 other;

Query Match 77.0%; Score 279.6; DB 20; Length 339;
 Best Local Similarity 93.9%; Pred. No. 7.3e-58;
 Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy     114 GAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db      90 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCCTA 149

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTCTGGCTCTGTACAACCAGCACAAACC 233
      |||
Db     150 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGCCCTGTACAACCAGCATAAACC 209

Qy     234 GGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db     210 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 269

Qy     294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 353
      |||||
Db     270 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 329

Qy     354 GTGCAGCTGA 363
      |||||
Db     330 GTGCAGCTGA 339

```

ID ABK84023 standard; cDNA; 650 BP.
XX
AC ABK84023;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #594.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US30821.
XX
PR 03-OCT-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
PS Claim 1; SEQ ID No 594; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen

DT 15-MAY-2002 (first entry)
XX
DE Kidney cancer related gene sequence SEQ ID NO:7155.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.

PR 01-NOV-2000; 2000US-245084P.

PA (AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;

DR WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -

PS Claim 1; SEQ ID 7155; 44pp; English.

CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.

SQ Sequence 650 BP; 108 A; 244 C; 179 G; 119 T; 0 other;

Best Local Similarity 93.9%; Pred. No. 8.2e-58;

Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Db 30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Db 90 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCTA 149

Db 150 CATTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCC 209

Db 210 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 269

```

Db      270  |||||CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAA 329
Qy      354  GTGCAGCTGA 363
Db      330  |||||GTGCAGCTGA 339

```

RESULT 13

ABZ35738

ID ABZ35738 standard; DNA; 1176 BP.

XX

AC ABZ35738;

XX

DT 07-FEB-2003 (first entry)

XX

DE Human TGF beta 1 polynucleotide SEQ ID NO 46.

XX

KW Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
 KW protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
 KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
 KW Hepatitis C virus; human papilloma virus; gene; ds.

XX

OS Homo sapiens.

XX

PN DE10100588-A1.

XX

PD 18-JUL-2002.

XX

PF 09-JAN-2001; 2001DE-1000588.

XX

PR 09-JAN-2001; 2001DE-1000588.

XX

PA (RIBO-) RIBOPHARMA AG.

XX

PI Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX

DR WPI; 2002-683450/74.

XX

PT Inhibiting expression of target genes, useful e.g. for treating tumors,
 PT by introducing into cells two double-stranded RNAs that are
 PT complementary to the target -

XX

PS Claim 13; Page 34-35; 100pp; German.

XX

CC The invention relates to inhibiting expression of a target gene in a cell
 CC by introducing at least two oligoribonucleotides (dsRNAI and II), both
 CC with a double-stranded (ds) structure of at most 49 sequential nucleotide
 CC pairs. At least part of one strand (S1, S2) of the ds structures in each
 CC of dsRNAI and II are complementary to regions in the target gene. The
 CC method uses antisense inhibition of gene expression using double stranded
 CC RNA inhibition (RNAi). The method is particularly used to treat tumours
 CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on
 CC humans, animals or plants). The method provides more effective inhibition
 CC of expression than known methods using a single dsRNA, even at very low
 CC concentrations. When dsRNA has at least one unpaired nucleotide at the
 CC end, stability (and thus effective concentration in the cell) is

CC improved and efficiency can be increased further by pretreating the cells
CC with interferon. The present sequence is that of a target DNA of the
CC invention.

XX

SQ Sequence 1176 BP; 239 A; 382 C; 353 G; 202 T; 0 other;

Query Match 77.0%; Score 279.6; DB 24; Length 1176;
Best Local Similarity 93.9%; Pred. No. 9.1e-58;
Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      867 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 926

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      ||||| || || |||||||||||||||||||||||||||||||||||||
Db      927 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGCCCCCTA 986

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAAACCC 233
      ||| ||||||| ||||| ||||||| ||||||| ||||||| ||||||| |||||
Db      987 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCC 1046

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db     1047 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 1106

Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 353
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db     1107 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 1166

Qy      354 GTGCAGCTGA 363
      |||||||||
Db     1167 GTGCAGCTGA 1176
```

RESULT 14

ABX09981

ID ABX09981 standard; DNA; 1176 BP.

XX

AC ABX09981;

XX

DT 23-JAN-2003 (first entry)

XX

DE Human TGFbeta1 DNA fragment SEQ ID 46.

XX

KW Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;
KW prion; inhibition; human; ds.

XX

OS Homo sapiens.

XX

PN DE10100587-C1.

XX

PD 21-NOV-2002.

XX

PF 09-JAN-2001; 2001DE-1000587.

XX

PR 09-JAN-2001; 2001DE-1000587.

```

XX      (RIBO-) RIBOPHARMA AG.
XX
PI      Kreutzer R,  Limmer S,  Rost S,  Hadwiger P;
XX
DR      WPI; 2002-742209/81.
XX
PT      Inhibiting expression of target genes, e.g. oncogenes, in cells, by
PT      introduction of complementary double-stranded oligoribonucleotide,
PT      after treating the cell with interferon -
XX
PS      Disclosure; Page 39-40; 98pp; German.
XX
CC      This invention describes a novel method for inhibiting expression of a
CC      target gene by introducing into the cell that contains the target gene
CC      at least one oligoribonucleotide (dsRNAI) that has a double-stranded
CC      (ds) structure of not more than 49 consecutive nucleotides (nt), where
CC      at least a segment of one strand of the ds structure is complementary
CC      with the target gene and the cells are treated with interferon before
CC      introduction of dsRNAI. The method is used to inhibit expression of
CC      target genes, particularly oncogenes, cytokine genes, Id (not defined)
CC      protein genes; developmental or prion genes, or genes expressed in
CC      pathogenic organisms (particularly plasmodia) or in viruses or viroids
CC      (pathogenic in humans, animals or plants). Treating the cells with
CC      interferon greatly increases the extent to which dsRNA can inhibit
CC      expression of the target genes, and the effect is even greater when dsRNA
CC      are modified to increase their stability. ABX09936-ABX10075 represent
CC      gene fragments used to illustrate the method of the invention.
XX
SQ      Sequence 1176 BP; 239 A; 382 C; 353 G; 202 T; 0 other;

Query Match          77.0%;  Score 279.6;  DB 24;  Length 1176;
Best Local Similarity 93.9%;  Pred. No. 9.1e-58;
Matches 291;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;

Qy      54  CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      867  CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 926

Qy      114  GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db      927  GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCCCTA 986

Qy      174  CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAACCC 233
      |||
Db      987  CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCC 1046

Qy      234  GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db      1047  GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTA 1106

Qy      294  CTACGTGGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTTCGTTCTGCAA 353
      |||||
Db      1107  CTACGTGGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTTCGTTCTGCAA 1166

Qy      354  GTGCAGCTGA 363
      |||||

```

Db 1167 GTGCAGCTGA 1176

RESULT 15

ABV78162

ID ABV78162 standard; DNA; 1176 BP.

XX

AC ABV78162;

XX

DT 15-NOV-2002 (first entry)

XX

DE Human TGF beta 1 DNA SEQ ID NO 46.

XX

KW RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic;
KW virucide; protozoacide; gene; ds.

XX

OS Homo sapiens.

XX

PN WO200255693-A2.

XX

PD 18-JUL-2002.

XX

PF 09-JAN-2002; 2002WO-EP00152.

XX

PR 09-JAN-2001; 2001DE-1000586.

PR 26-OCT-2001; 2001DE-1055280.

PR 29-NOV-2001; 2001DE-1058411.

PR 07-DEC-2001; 2001DE-1060151.

XX

PA (RIBO-) RIBOPHARMA AG.

XX

PI Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX

DR WPI; 2002-590671/63.

XX

PT Inhibiting expression of target gene, useful e.g. for inhibiting
PT oncogenes, by administering double-stranded RNA complementary to the
PT target and having an overhang -

XX

PS Claim 10; Page 138; 203pp; German.

XX

CC The invention relates to inhibiting expression of a target gene (I) in a
CC cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded
CC structure of at most 49 consecutive bases. At least part of one strand
CC (as1) of dsRNA1 is complementary to (I) and at least one end of dsRNA1
CC has an overhang of 1-4 nucleotides. The method is used to inhibit the
CC expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
CC in humans, also genes in Plasmodium or in viruses or viroids that are
CC pathogenic for humans, animals or plants. Introducing an overhang into
CC dsRNA greatly increases effectiveness for inhibiting gene expression,
CC both in vivo and in vitro and also increases stability and thus the
CC effective concentration inside the cell. The present sequence is that of
CC a gene related to the invention.

XX

SQ Sequence 1176 BP; 239 A; 382 C; 353 G; 202 T; 0 other;

Query Match

77.0%; Score 279.6; DB 24; Length 1176;

Best Local Similarity 93.9%; Pred. No. 9.1e-58;
Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      867 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 926

Qy     114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db     927 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTA 986

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAACCC 233
      |||||
Db     987 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCC 1046

Qy     234 GGGCGCGTCGGCGGCGCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db    1047 GGGCGCCTCGGCGGCGCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 1106

Qy     294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAA 353
      |||||
Db    1107 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTCTCTTGCAA 1166

Qy     354 GTGCAGCTGA 363
      |||||
Db    1167 GTGCAGCTGA 1176
```

Search completed: October 27, 2003, 19:11:03
Job time : 110.057 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 18:35:27 ; Search time 107.894 Seconds
(without alignments)
9022.658 Million cell updates/sec

Title: US-10-017-372E-14
Perfect score: 363
Sequence: 1 gccctggataccaactactg.....gttcctgcaagtgcagctga 363

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%						Description
	No.	Score	Query Match	Length	DB ID		
	1	281.2	77.5	489	11	US-09-911-904-167	Sequence 167, App
	2	279.6	77.0	339	10	US-09-813-271B-1	Sequence 1, Appli
	3	279.6	77.0	1821	14	US-10-087-268-1	Sequence 1, Appli
	4	279.6	77.0	1821	14	US-10-087-268-4	Sequence 4, Appli
	5	279.6	77.0	2742	14	US-10-037-270-220	Sequence 220, App
	6	279.6	77.0	2745	11	US-09-948-002-28	Sequence 28, Appl
	7	257.2	70.9	2094	11	US-09-948-002-1	Sequence 1, Appli
	8	255.6	70.4	1585	11	US-09-948-002-27	Sequence 27, Appl
	9	198.2	54.6	336	10	US-09-813-271B-7	Sequence 7, Appli
	10	191.4	52.7	339	10	US-09-813-271B-5	Sequence 5, Appli
	11	191.4	52.7	2574	11	US-09-906-158-3	Sequence 3, Appli
	12	191.4	52.7	2574	13	US-10-028-158-20	Sequence 20, Appl
	13	188.2	51.8	2879	11	US-09-906-158-10	Sequence 10, Appl
	14	188.2	51.8	4382	12	US-09-957-458B-9	Sequence 9, Appli
	15	169.6	46.7	336	10	US-09-813-271B-9	Sequence 9, Appli
	16	164.4	45.3	336	10	US-09-813-271B-11	Sequence 11, Appl
	17	155	42.7	4267	11	US-09-948-002-47	Sequence 47, Appl
	18	147	40.5	339	10	US-09-813-271B-3	Sequence 3, Appli
	19	147	40.5	2570	12	US-09-960-706-663	Sequence 663, App
	20	147	40.5	2912	13	US-10-044-090-323	Sequence 323, App
c	21	131	36.1	597	9	US-09-864-761-15319	Sequence 15319, A
c	22	130.2	35.9	206	10	US-09-833-381-577	Sequence 577, App
c	23	128.6	35.4	154	9	US-09-864-761-31841	Sequence 31841, A
	24	103.2	28.4	851	13	US-10-027-632-152938	Sequence 152938,
c	25	98.8	27.2	224	10	US-09-833-381-73	Sequence 73, Appl
	26	98.4	27.1	29000	11	US-09-906-158-17	Sequence 17, Appl
	27	73.2	20.2	450	11	US-09-918-995-6145	Sequence 6145, Ap

	28	69	19.0	584	12	US-10-029-386-9758	Sequence 9758, Ap
	29	68.2	18.8	176	12	US-10-029-386-23458	Sequence 23458, A
c	30	66.2	18.2	176	10	US-09-833-381-571	Sequence 571, App
	31	57.4	15.8	486	11	US-09-918-995-25641	Sequence 25641, A
	32	46.6	12.8	412	10	US-09-841-730-28	Sequence 28, Appl
	33	46.6	12.8	476	10	US-09-841-730-26	Sequence 26, Appl
	34	45.4	12.5	573	14	US-10-156-761-3931	Sequence 3931, Ap
	35	45.4	12.5	9025608	14	US-10-156-761-1	Sequence 1, Appli
	36	45	12.4	630	9	US-09-454-540-3	Sequence 3, Appli
	37	45	12.4	630	10	US-09-859-894A-3	Sequence 3, Appli
	38	45	12.4	630	12	US-09-871-604-3	Sequence 3, Appli
	39	44.2	12.2	557	9	US-09-765-527-248	Sequence 248, App
	40	44.2	12.2	557	9	US-09-760-397-1	Sequence 1, Appli
	41	44.2	12.2	557	12	US-10-324-182-1	Sequence 1, Appli
	42	44.2	12.2	610	9	US-09-760-397-7	Sequence 7, Appli
	43	44.2	12.2	610	12	US-10-324-182-7	Sequence 7, Appli
	44	44.2	12.2	613	9	US-09-765-527-256	Sequence 256, App
	45	44.2	12.2	613	9	US-09-765-527-260	Sequence 260, App

ALIGNMENTS

RESULT 1

```

US-09-911-904-167
; Sequence 167, Application US/09911904
; Publication No. US20030096234A1
; GENERAL INFORMATION:
; APPLICANT: Farr, Spencer B.
; APPLICANT: Pickett, Gavin G.
; APPLICANT: Neft, Robin Eileen
; APPLICANT: Dunn, II, Robert Thomas
; TITLE OF INVENTION: CANINE TOXICITY GENES
; FILE REFERENCE: 400742000200
; CURRENT APPLICATION NUMBER: US/09/911,904
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/220,057
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 167
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(489)
; OTHER INFORMATION: n = A, T, C or G
US-09-911-904-167

```

```

Query Match          77.5%;  Score 281.2;  DB 11;  Length 489;
Best Local Similarity 94.2%;  Pred. No. 9e-74;
Matches 292;  Conservative 0;  Mismatches 18;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
        || |||||
Db      101 CACGGAGAAGAACTGCTGCGTCCGCGCAGCTCTACATTGACTTCCGCAAGGATCTGGGCTG 160

```



```

;          TELEPHONE: (908) 522-6940
;          TELEFAX: (908) 522-6955
;  INFORMATION FOR SEQ ID NO: 1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 339 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;    MOLECULE TYPE: cDNA to mRNA
;    HYPOTHETICAL: NO
;    IMMEDIATE SOURCE:
;      CLONE: E. coli LC137/pPLMu.hTGF-beta1 (DSM 5656)
;    FEATURE:
;      NAME/KEY: CDS
;      LOCATION: 1..336
;      OTHER INFORMATION: /product= "human TGF-beta1"
;    SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-813-271B-1

```

```

Query Match          77.0%;  Score 279.6;  DB 10;  Length 339;
Best Local Similarity 93.9%;  Pred. No. 2.5e-73;
Matches 291;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy     114 GAAGTGGATTTCATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db      90 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCTA 149

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAACCC 233
      |||||
Db     150 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCC 209

Qy     234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db     210 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 269

Qy     294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 353
      |||||
Db     270 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 329

Qy     354 GTGCAGCTGA 363
      |||||
Db     330 GTGCAGCTGA 339

```

RESULT 3

US-10-087-268-1

```

; Sequence 1, Application US/10087268
; Publication No. US20030119010A1
; GENERAL INFORMATION:
; APPLICANT: Jonsonn, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or
condition

```



```

; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1821
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (1)..(511)
;   OTHER INFORMATION:
;   NAME/KEY: CDS
;   LOCATION: (512)..(1684)
;   OTHER INFORMATION:
;   NAME/KEY: sig_peptide
;   LOCATION: (512)..(598)
;   OTHER INFORMATION:
;   NAME/KEY: 3'UTR
;   LOCATION: (1685)..(1821)
;   OTHER INFORMATION:
US-10-087-268-1

```

```

Query Match          77.0%;  Score 279.6;  DB 14;  Length 1821;
Best Local Similarity 93.9%;  Pred. No. 3.5e-73;
Matches 291;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1375 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 1434

Qy     114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      ||||||||| || || |||||||||||||||||||||||||||||||||||
Db     1435 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTA 1494

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTTGGCTCTGTACAACCAGCACAAACCC 233
      ||| ||||||||| ||||| ||||||||| ||||||||||||||| |||||
Db     1495 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGGCCCTGTACAACCAGCATAAACCC 1554

Qy     234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      ||||| ||||||||||||||||||||||||||||||||||| |||||||
Db     1555 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 1614

Qy     294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
      ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db     1615 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 1674

Qy     354 GTGCAGCTGA 363
      |||||||||
Db     1675 GTGCAGCTGA 1684

```

```

RESULT 4
US-10-087-268-4
; Sequence 4, Application US/10087268
; Publication No. US20030119010A1

```

```
; GENERAL INFORMATION:
; APPLICANT: Jonsonn, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or
condition
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(511)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (512)..(1684)
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (512)..(598)
; OTHER INFORMATION:
; NAME/KEY: 3'UTR
; LOCATION: (1685)..(1821)
; OTHER INFORMATION:
US-10-087-268-4
```

```
Query Match          77.0%; Score 279.6; DB 14; Length 1821;
Best Local Similarity 93.9%; Pred. No. 3.5e-73;
Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
```

```
Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      1375 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 1434

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db      1435 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTA 1494

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTCTGGCTCTGTACAACCAGCACAAACC 233
      |||
Db      1495 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTCTGGCCCTGTACAACCAGCATAAACC 1554

Qy      234 GGGCGCGTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db      1555 GGGCGCCTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 1614

Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 353
      |||||
Db      1615 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 1674

Qy      354 GTGCAGCTGA 363
      |||||
Db      1675 GTGCAGCTGA 1684
```

RESULT 5

US-10-037-270-220

```
; Sequence 220, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 220
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (842)..(2014)
US-10-037-270-220
```

```
Query Match          77.0%;  Score 279.6;  DB 14;  Length 2742;
Best Local Similarity 93.9%;  Pred. No. 3.9e-73;
Matches 291;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;
```

```
Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1705 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 1764

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      ||||| || || |||||||||||||||||||||||||||||||||||
Db      1765 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGCCCCTA 1824

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAAACC 233
      ||| ||||||| ||||| ||||||| ||||||||||||||| |||||||
```

```

Db      1825 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAAACC 1884
Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1885 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCCGCTGCCCATCGTGTA 1944
Qy      294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1945 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAA 2004
Qy      354 GTGCAGCTGA 363
        ||||| |||||
Db      2005 GTGCAGCTGA 2014

```

RESULT 6

```

US-09-948-002-28
; Sequence 28, Application US/09948002
; Publication No. US20030050265A1
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Susan F. Murray
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH
; TITLE OF INVENTION: FACTOR BETA EXPRESSION
; FILE REFERENCE: ISPH-0607
; CURRENT APPLICATION NUMBER: US/09/948,002
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/661,753
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/154,546
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 71
; SEQ ID NO 28
; LENGTH: 2745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (842)...(2017)
US-09-948-002-28

```

```

Query Match          77.0%;  Score 279.6;  DB 11;  Length 2745;
Best Local Similarity 93.9%;  Pred. No. 3.9e-73;
Matches 291;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
        || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1708 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 1767
Qy      114 GAAGTGGATTGATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1768 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 1827
Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAACCC 233
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1828 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAAACC 1887

```

```

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||
Db      1888 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 1947

Qy      294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTGCAA 353
      |||
Db      1948 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTCTCTGCAA 2007

Qy      354 GTGCAGCTGA 363
      |||
Db      2008 GTGCAGCTGA 2017

```

RESULT 7

US-09-948-002-1

```

; Sequence 1, Application US/09948002
; Publication No. US20030050265A1
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Susan F. Murray
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH
; TITLE OF INVENTION: FACTOR BETA EXPRESSION
; FILE REFERENCE: ISPH-0607
; CURRENT APPLICATION NUMBER: US/09/948,002
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/661,753
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/154,546
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 71
; SEQ ID NO 1
; LENGTH: 2094
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (868)...(2040)

```

US-09-948-002-1

```

Query Match          70.9%; Score 257.2; DB 11; Length 2094;
Best Local Similarity 89.4%; Pred. No. 1.6e-66;
Matches 277; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      ||
Db      1731 CACAGAGAAGAACTGCTGTGTGCGGCAGCTGTACATTGACTTTAGGAAGGACCTGGGTTG 1790

Qy      114 GAAGTGGATTTCATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||
Db      1791 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGTCTGGGACCCTGCCCTA 1850

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTCTGGCTCTGTACAACCAGCACAAACC 233
      |||
Db      1851 TATTTGGAGCCTGGACACACAGTACAGCAAGGTCTCTGCCCTCTACAACCAACACAACCC 1910

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||

```

```

Db      1911 GGGCGCTTCGGCGTCAACGTGCTGCGTGCCGCAGGCTTTGGAGCCACTGCCCATCGTCTA 1970
Qy      294 CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
      |||||
Db      1971 CTACGTGGGTCGCAAGCCCAAGGTGGAGCAGTTGTCCAACATGATTGTGCGCTCCTGCAA 2030
Qy      354 GTGCAGCTGA 363
      |||||
Db      2031 GTGCAGCTGA 2040

```

RESULT 8

US-09-948-002-27

; Sequence 27, Application US/09948002

; Publication No. US20030050265A1

; GENERAL INFORMATION:

; APPLICANT: Nicholas M. Dean

APPLICANT: Susan F. Murray

TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH

; TITLE OF INVENTION: FACTOR BETA EXPRESSION

FILE REFERENCE: ISPH-0607

; CURRENT APPLICATION NUMBER: US/09/948,002

; CURRENT FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: 09/661,753

; PRIOR FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: 60/154,546

;
; PRIOR FILING DATE: 1999-09-17

; NUMBER OF SEQ ID NOS: 71

; SEQ ID NO 27

; LENGTH: 1585

; TYPE: DNA

ORGANISM: *Rattus norvegicus*

FEATURE:

; NAME/KEY: CDS

LOCATION: (413)...(1585)

US-09-948-002-27

Query Match 70.4%; Score 255.6; DB 11; Length 1585;

Best Local Similarity 89.0%; Pred. No. 4.6e-66;

Matches 276; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

[illegible]

```

Qy      294 CTACGTGGGCGCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 353
          |||
Db      1516 CTACGTGGGTCGCAAGCCCCAAGGTGGAGCAGTTGTCCAACATGATCGTGCCTTCCTGCAA 1575

Qy      354 GTGCAGCTGA 363
          |||
Db      1576 GTGCAGCTGA 1585

```

RESULT 9

US-09-813-271B-7

; Sequence 7, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

; APPLICANT:

; (A) Nico Cerletti

; TITLE OF INVENTION: New process for the production of
; biologically active protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. US20020115834A1artis Patent Department

; STREET: 564 Morris Avenue

; CITY: Summit

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07901

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/813,271B

; FILING DATE: 20-Mar-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP95/02719

; FILING DATE: 12-Jul-95

; APPLICATION NUMBER: EPO 94810439.3

; FILING DATE: 25-Jul-94

; ATTORNEY/AGENT INFORMATION:

; NAME: Pfeiffer, Hesna J. .

; REGISTRATION NUMBER: 22640

; REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (908) 522-6940

; TELEFAX: (908) 522-6955

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 336 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; DESCRIPTION: /desc = "recombinant hybrid DNA of

; IMMEDIATE SOURCE:

; CLONE: E. coli LC137/pPLMu.TGF-beta1(44/45)beta3

; FEATURE:

```

;          NAME/KEY:  mat_peptide
;          LOCATION:  1..132
;          OTHER INFORMATION:  /product= "N-terminal 44 amino
;          acids of human TGF-beta1"
;  FEATURE:
;          NAME/KEY:  mat_peptide
;          LOCATION:  133..336
;          OTHER INFORMATION:  /product= "C-terminal 68 amino
;          acids of human TGF-beta3"
;  FEATURE:
;          NAME/KEY:  CDS
;          LOCATION:  1..336
;          OTHER INFORMATION:  /product= "hybrid TGF-beta named
;          TGF-beta1-3"
;  SEQUENCE DESCRIPTION:  SEQ ID NO: 7:
US-09-813-271B-7

```

```

Query Match          54.6%;  Score 198.2;  DB 10;  Length 336;
Best Local Similarity 77.9%;  Pred. No. 3.7e-49;
Matches 239;  Conservative 0;  Mismatches 68;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      30 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 89

Qy     114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      || |||||
Db      90 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCTCAGGCCCTTGCCCATA 149

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGGCTCTGTACAACCAGCACAAACCC 233
      || || || || |||||
Db     150 CCTCCGCACTGCAGACACAACCCACAGCACGGTGCTGGGACTGTACAACACTCTGAACCC 209

Qy     234 GGGCGCGTCGGCGGCGCGGCTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      | || || || || |||||
Db     210 TGAAGCATCTGCCTCGCCTTGCTGCGTGCCCCAGGACCTGGAGCCCCTGACCATCCTGTA 269

Qy     294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAA 353
      ||| || || || |||||
Db     270 CTATGTTGGGAGGACCCCCAAAGTGGAGCAGCTCTCCAACATGGTGGTGAAGTCTTGTA 329

Qy     354 GTGCAGC 360
      || |||
Db     330 ATGTAGC 336

```

RESULT 10

US-09-813-271B-5

```

; Sequence 5, Application US/09813271B
; Patent No. US20020115834A1
;  GENERAL INFORMATION:
;  APPLICANT:
;
;          (A) Nico Cerletti
;  TITLE OF INVENTION:  New process for the production of
;                      biologically active protein
;  NUMBER OF SEQUENCES: 13
;  CORRESPONDENCE ADDRESS:

```



```

;      ADDRESSEE: No. US20020115834Alartis Patent Department
;      STREET: 564 Morris Avenue
;      CITY: Summit
;      STATE: New Jersey
;      COUNTRY: USA
;      ZIP: 07901
;
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/813,271B
;      FILING DATE: 20-Mar-2001
;
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: PCT/EP95/02719
;      FILING DATE: 12-Jul-95
;      APPLICATION NUMBER: EPO 94810439.3
;      FILING DATE: 25-Jul-94
;
;      ATTORNEY/AGENT INFORMATION:
;      NAME: Pfeiffer, Hesna J. .
;      REGISTRATION NUMBER: 22640
;      REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN
;
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (908) 522-6940
;      TELEFAX: (908) 522-6955
;
;      INFORMATION FOR SEQ ID NO: 5:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 339 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;      MOLECULE TYPE: cDNA to mRNA
;      IMMEDIATE SOURCE:
;      CLONE: E. coli LC137/pPLMu.hTGF-beta3 (DSM 5658)
;      FEATURE:
;      NAME/KEY: CDS
;      LOCATION: 1..336
;      OTHER INFORMATION: /product= "human TGF-beta3"
;      SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-813-271B-5

```

```

Query Match          52.7%;  Score 191.4;  DB 10;  Length 339;
Best Local Similarity 75.7%;  Pred. No. 3.8e-47;
Matches 237;  Conservative 0;  Mismatches 76;  Indels 0;  Gaps 0;

```

```

Qy      51  CGACAAGGAGAAGAACTGCTGCGTGC GGCAGCTCTACATTGACTTCCGGAAGGACCTGGG 110
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      27  CAACTTGGAGGAGAAGAACTGCTGTGTGCGCCCCCTCTACATTGACTTCCGACAGGATCTGGG 86

Qy     111  CTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCC 170
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      87  CTGGAAGTGGGTCCATGAACCTAAGGGCTACTATGCCAATTCTGCTCAGGCCCTTGCCC 146

Qy     171  CTACATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAA 230
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     147  ATACCTCCGCAGTGCAGACACAACCCACAGCACGGTGCTGGGACTGTACAACACTCTGAA 206

```

```

Qy      231 CCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGT 290
      ||| | || || || || || || || || || || || || || || || || || || || ||
Db      207 CCCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCCCAGGACCTGGAGCCCCTGACCATCCT 266

Qy      291 GTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTG 350
      ||||| || || || || || || || || || || || || || || || || || || || ||
Db      267 GTACTATGTTGGGAGGACCCCCAAAGTGGAGCAGCTCTCCAACATGGTGGTGAAGTCTTG 326

Qy      351 CAAGTGCAGCTGA 363
      || || || || || || || ||
Db      327 TAAATGTAGCTGA 339

```

RESULT 11

```

US-09-906-158-3
; Sequence 3, Application US/09906158
; Publication No. US20030078217A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH FACTOR-BETA 3
EXPRESSION
; FILE REFERENCE: RTS-0257
; CURRENT APPLICATION NUMBER: US/09/906,158
; CURRENT FILING DATE: 2001-07-14
; NUMBER OF SEQ ID NOS: 168
; SEQ ID NO 3
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (254)...(1492)
US-09-906-158-3

```

```

Query Match          52.7%; Score 191.4; DB 11; Length 2574;
Best Local Similarity 75.7%; Pred. No. 5.9e-47;
Matches 237; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```

```

Qy      51 CGACAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGG 110
      ||| ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db      1180 CAACTTGGAGGAGAAGCTGCTGTGTGCGCCCCCTCTACATTGACTTCCGACAGGATCTGGG 1239

Qy      111 CTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCC 170
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||
Db      1240 CTGGAAGTGGGTCCATGAACCTAAGGGCTACTATGCCAATTCTGCTCAGGCCCTTGCCC 1299

Qy      171 CTACATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAA 230
      ||| || | || ||||| ||| || ||| |||| ||||| ||||| || ||
Db      1300 ATACCTCCGCAGTGCAGACACAACCCACAGCACGGTGCTGGGACTGTACAACACTCTGAA 1359

Qy      231 CCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGT 290
      ||| | || || || || || || || || || || || || || || || || || || || ||
Db      1360 CCCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCCCAGGACCTGGAGCCCCTGACCATCCT 1419

Qy      291 GTACTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTG 350

```

```

          ||||| || || | | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1420 GTACTATGTTGGGAGGACCCCAAGTGAGCAGCTCTCCAACATGGTGGTGAAGTCTTG 1479

Qy      351 CAAGTGAGCTGA 363
          || || |||||
Db      1480 TAAATGTAGCTGA 1492

```

RESULT 12

US-10-028-158-20

```

; Sequence 20, Application US/10028158
; Publication No. US20020110833A1
; GENERAL INFORMATION:
; APPLICANT: Caniggia, Isabella
; APPLICANT: Post, Martin
; APPLICANT: Lye, Stephen
; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF
; TITLE OF INVENTION: TROPHOBLAST
; FILE REFERENCE: 11757.38USWO
; CURRENT APPLICATION NUMBER: US/10/028,158
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US/09/380,662
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: PCT/CA98/00180
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (254)..(1492)
US-10-028-158-20

```

```

Query Match          52.7%; Score 191.4; DB 13; Length 2574;
Best Local Similarity 75.7%; Pred. No. 5.9e-47;
Matches 237; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```

```

Qy      51 CGACAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGG 110
          | || | |||| | ||||| ||||| | | ||||| ||||| ||||| |||||
Db      1180 CAACTTGGAGGAGAAGAACTGCTGTGTGCGCCCCCTCTACATTGACTTCCGACAGGATCTGGG 1239

Qy      111 CTGGAAGTGGATTTCATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCC 170
          ||||| ||||| | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1240 CTGGAAGTGGGTCCATGAACCTAAGGGCTACTATGCCAACTTCTGCTCAGGCCCTTGCCC 1299

Qy      171 CTACATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAA 230
          ||| || | || | ||||| ||| || ||| |||| | ||||| ||||| |||
Db      1300 ATACCTCCGCAGTGCAGACACAACCCACAGCACGGTGCTGGGACTGTACAACACTCTGAA 1359

Qy      231 CCCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGT 290
          ||| | || || || || |||| | ||||| ||||| ||||| ||||| |||||
Db      1360 CCCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCCCAGGACCTGGAGCCCCTGACCATCCT 1419

```

```

Qy      291 GTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTG 350
      ||||| || || | | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1420 GTACTATGTTGGGAGGACCCCCAAAGTGGAGCAGCTCTCCAACATGGTGGTGAAGTCTTG 1479

Qy      351 CAAGTGCAGCTGA 363
      || || |||||
Db      1480 TAAATGTAGCTGA 1492

```

RESULT 13

US-09-906-158-10

; Sequence 10, Application US/09906158

; Publication No. US20030078217A1

; GENERAL INFORMATION:

; APPLICANT: Brett P. Monia

; APPLICANT: Susan M. Freier

; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH FACTOR-BETA 3
EXPRESSION

; FILE REFERENCE: RTS-0257

; CURRENT APPLICATION NUMBER: US/09/906,158

; CURRENT FILING DATE: 2001-07-14

; NUMBER OF SEQ ID NOS: 168

; SEQ ID NO 10

; LENGTH: 2879

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (611)...(1843)

US-09-906-158-10

```

Query Match          51.8%; Score 188.2; DB 11; Length 2879;
Best Local Similarity 75.1%; Pred. No. 5.4e-46;
Matches 235; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```

```

Qy      51 CGACAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGG 110
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1531 CAACCTGGAGGAGAAGAACTGCTGTGTACGCCCCCTTTATATTGACTTCCGGCAGGATCTAGG 1590

Qy      111 CTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCC 170
      ||||| ||| | || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1591 CTGGAAATGGGTCCACGAACCTAAGGGTTACTATGCCAACTTCTGCTCAGGCCCTTGCCC 1650

Qy      171 CTACATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAA 230
      ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1651 ATACCTCCGCAGCGCAGACACAACCCATAGCACGGTGCTTGGACTATACAACACCCTGAA 1710

Qy      231 CCCGGGCGCGTCCGGCGGCGCGTGTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGT 290
      ||| | ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1711 CCCAGAGGCGTCTGCCTCGCCATGCTGCGTCCCCCAGGACCTGGAGCCCCTGACCATCTT 1770

Qy      291 GTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTG 350
      ||||| ||||| || | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1771 GTACTATGTGGGCAGAACCCCAAGGTGGAGCAGCTGTCCAACATGGTGGTGAAGTCGTG 1830

Qy      351 CAAGTGCAGCTGA 363

```


RESULT 15

US-09-813-271B-9

; Sequence 9, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

; APPLICANT:

; (A) Nico Cerletti

; TITLE OF INVENTION: New process for the production of
; biologically active protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. US20020115834A1artis Patent Department

; STREET: 564 Morris Avenue

; CITY: Summit

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07901

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/813,271B

; FILING DATE: 20-Mar-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP95/02719

; FILING DATE: 12-Jul-95

; APPLICATION NUMBER: EPO 94810439.3

; FILING DATE: 25-Jul-94

; ATTORNEY/AGENT INFORMATION:

; NAME: Pfeiffer, Hesna J. .

; REGISTRATION NUMBER: 22640

; REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (908) 522-6940

; TELEFAX: (908) 522-6955

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 336 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: other nucleic acid

; DESCRIPTION: /desc = "recombinant hybrid DNA

; IMMEDIATE SOURCE:

; CLONE: E. coli LC137/pPLMu.TGF-beta2(44/45)beta3

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: 1..132

; OTHER INFORMATION: /product= "N-terminal 44 amino
; acids of human TGF-beta2"

; FEATURE:

; NAME/KEY: mat_peptide

; LOCATION: 133..336

```

;           OTHER INFORMATION: /product= "C-terminal 68 amino
;           acids of human TGF-beta3"
;   FEATURE:
;           NAME/KEY:   CDS
;           LOCATION:   1..336
;           OTHER INFORMATION: /product= "hybrid TGF-beta2-3"
;   SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-813-271B-9

```

Query Match 46.7%; Score 169.6; DB 10; Length 336;
Best Local Similarity 72.4%; Pred. No. 1.1e-40;
Matches 220; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy	57	GGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAA	116
Db	33	GCAGGATAATTGCTGCCTACGTCCA CTTTACATTGATTTCAAGAGGGATCTAGGGTGGAA	92
Qy	117	GTGGATT CATGAACCCAAGGGCTACC ATGCCAATTTCTGCCTGGGGCCCTGTCCCTACAT	176
Db	93	ATGGATACACGAACCCAAAGGGTACAATGCCA ACTTCTGCTCAGGCCCTTGCCCATACCT	152
Qy	177	CTGGAGCCTAGACACTCAGTACAACAAGGTCTCTGGCTCTGTACAACCAGCACAA CCCGGG	236
Db	153	CCGCAGTGCAGACACAACCCACAGCACGGTGTCTGGGACTGTACA A CACTCTGAAC CCTGA	212
Qy	237	CGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTA	296
Db	213	AGCATCTGCCTCGCCTTGCTGCGTGCCCCAGGACCTGGAGCCCCTGACCATCCTGTACTA	272
Qy	297	CGTGGGCCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGT GCGTTCTTGCAAGTG	356
Db	273	TGTTGGGAGGACCCCCAAAGTGGAGCAGCTCTCCAACATGGTGGTGAAGTCTTGTA AATG	332
Qy	357	CAGC	360
Db	333	TAGC	336

Search completed: October 28, 2003, 09:04:19
Job time : 113.037 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

```
Run on:      October 27, 2003, 10:47:27 ; Search time 900.804 Seconds
              (without alignments)
              9794.056 Million cell updates/sec
```

Title: US-10-017-372E-14
Perfect score: 363
Sequence: 1 gccctggataccaactactg.....gttcctgcaagtgcagctga 363

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%		Query		ID	Description
	No.	Score	Match	Length	DB			
c	1	279.8	77.1	546	9	AI760533	AI760533	wh88b09.x
c	2	279.8	77.1	954	13	BX392116	BX392116	BX392116
	3	279.6	77.0	396	12	BI868266	BI868266	603392221
	4	279.6	77.0	400	12	BI820759	BI820759	603034355
	5	279.6	77.0	431	10	BG180040	BG180040	602329596

	6	279.6	77.0	461	12	BM740537	BM740537 K-EST0011
c	7	279.6	77.0	475	13	BU632941	BU632941 UI-H-DF0-
c	8	279.6	77.0	480	9	AI089904	AI089904 qa16a08.x
	9	279.6	77.0	505	10	AW958056	AW958056 EST370126
	10	279.6	77.0	536	13	BQ222205	BQ222205 AGENCOURT
c	11	279.6	77.0	538	12	BM670402	BM670402 UI-E-DW1-
	12	279.6	77.0	540	12	BI834732	BI834732 603090326
c	13	279.6	77.0	587	12	BQ016027	BQ016027 UI-H-DT1-
c	14	279.6	77.0	593	12	BQ003136	BQ003136 UI-H-EI1-
c	15	279.6	77.0	598	9	AI951831	AI951831 wx38b08.x
	16	279.6	77.0	599	12	BI905963	BI905963 603062849
c	17	279.6	77.0	600	10	BF726995	BF726995 by15c03.y
c	18	279.6	77.0	616	12	BM674987	BM674987 UI-E-EJ0-
c	19	279.6	77.0	625	14	CA426732	CA426732 UI-H-FE1-
c	20	279.6	77.0	649	14	CA423539	CA423539 UI-H-FE1-
c	21	279.6	77.0	654	12	BM681750	BM681750 UI-E-EJ0-
c	22	279.6	77.0	662	13	BQ576225	BQ576225 UI-H-EZ1-
c	23	279.6	77.0	674	14	CA503085	CA503085 UI-CF-FN0
c	24	279.6	77.0	676	13	BU737545	BU737545 UI-E-DW1-
c	25	279.6	77.0	677	14	CA432581	CA432581 UI-H-FL1-
c	26	279.6	77.0	678	12	BQ021537	BQ021537 UI-H-DH1-
c	27	279.6	77.0	690	13	BU625377	BU625377 UI-H-FG1-
c	28	279.6	77.0	695	14	CA425775	CA425775 UI-H-FE1-
c	29	279.6	77.0	697	13	BU633535	BU633535 UI-H-FL1-
c	30	279.6	77.0	697	14	CA426391	CA426391 UI-H-FE1-
c	31	279.6	77.0	700	13	BU633922	BU633922 UI-H-FL1-
c	32	279.6	77.0	722	13	BU632686	BU632686 UI-H-FE1-
c	33	279.6	77.0	757	14	CA309731	CA309731 UI-H-FT1-
c	34	279.6	77.0	888	13	BX434425	BX434425 BX434425
c	35	279.6	77.0	928	13	BX412434	BX412434 BX412434
c	36	279.6	77.0	956	12	BI084718	BI084718 602869722
	37	279.6	77.0	1181	12	BM805377	BM805377 AGENCOURT
	38	278.6	76.7	427	12	BM738551	BM738551 K-EST0006
	39	278.6	76.7	451	12	BM717493	BM717493 UI-E-EJ0-
	40	278.6	76.7	1093	12	BQ054305	BQ054305 AGENCOURT
c	41	278.6	76.7	1201	13	BX355682	BX355682 BX355682
c	42	278	76.6	474	10	BE645704	BE645704 7e74h10.x
c	43	278	76.6	583	13	BU625902	BU625902 UI-H-FG1-
c	44	278	76.6	655	13	BU624264	BU624264 UI-H-FG1-
c	45	278	76.6	659	14	CA423193	CA423193 UI-H-FE1-

ALIGNMENTS

RESULT 1

AI760533/c

LOCUS AI760533 546 bp mRNA linear EST 20-DEC-1999
 DEFINITION wh88b09.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2387801 3' similar to gb:X02812_cds1 TRANSFORMING GROWTH FACTOR BETA 1 PRECURSOR (HUMAN); contains element TAR1 repetitive element ;, mRNA sequence.

ACCESSION AI760533

VERSION AI760533.1 GI:5176200

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 546)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever, M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 714 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 465.

FEATURES Location/Qualifiers

source 1..546
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2387801"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_CLL1"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGATTGCTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 86 a 157 c 204 g 98 t 1 others

ORIGIN

Query Match 77.1%; Score 279.8; DB 9; Length 546;
Best Local Similarity 89.1%; Pred. No. 1.5e-60;
Matches 302; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

Qy      25 AGCTCCACGGACTACAAGGATGACGACGACAAGGAGAAGAACTGCTGCGTGCGGCAGCTC 84
      ||| |  || || |  |  |  || ||||| ||||| ||||| ||||| |||||
Db      476 AGCCCTGGACACAACCTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTG 417

Qy      85 TACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCAT 144
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      416 TACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCACGAGCCCAAGGGCTACCAT 357

Qy     145 GCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAACAAG 204
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     356 GCCAATTTCTGCCTGGGGCCCTGTCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAG 297

```

```

Qy      205 GTCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCG 264
          |||
Db      296 GTCCTGGCCCTGTACAACCAGCATAACCGGGCGCCTCGGCGGGCGCCGTGCTGCGTGCCG 237

Qy      265 CAGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGCGCAAGCCCAAGGTGGAGCAG 324
          |||
Db      236 CAGGCGCTGGAGCCGCTGCCCATCGTGTAACGTGGGCGCAAGCCCAAGGTGGAGCAG 177

Qy      325 CTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 363
          |||
Db      176 CTGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGA 138

```

RESULT 2

BX392116/c

LOCUS BX392116 954 bp mRNA linear EST 13-MAY-2003

DEFINITION BX392116 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DC005YF22 3-PRIME, mRNA sequence.

ACCESSION BX392116

VERSION BX392116.1 GI:30611808

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 954)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see

<http://www.genoscope.cns.fr/>

[cgi-bin/cluster.cgi?seq=CS0BAK013CE06NM1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0BAK013CE06NM1&cluster=9160.r). Contact :

Feng Liang Email : fliang@lifetech.com URL :

<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0BAK013CE06NM1.

FEATURES Location/Qualifiers

source

1. .954

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DC005YF22"

/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 187 a 258 c 307 g 202 t

ORIGIN

Query Match

77.1%; Score 279.8; DB 13; Length 954;


```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5402212"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

BASE COUNT 75 a 135 c 121 g 65 t
ORIGIN

Query Match 77.0%; Score 279.6; DB 12; Length 396;
Best Local Similarity 93.9%; Pred. No. 1.5e-60;
Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      70 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 129

Qy     114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db     130 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTA 189

Qy     174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTTGGCTCTGTACAACCAGCACAAACCC 233
      ||| |||||
Db     190 CATTTGGAGCCTGGACACGAGTACAGCAAGGTCTTGGCCCTGTACAACCAGCATAACCC 249

Qy     234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db     250 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCGCTGCCCATCGTGTA 309

Qy     294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTTGCAA 353
      |||||
Db     310 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTTGCAA 369

Qy     354 GTGCAGCTGA 363
      |||||
Db     370 GTGCAGCTGA 379

```

RESULT 4
BI820759

LOCUS BI820759 400 bp mRNA linear EST 04-OCT-2001
DEFINITION 603034355F1 NIH_MGC 115 Homo sapiens cDNA clone IMAGE:5175732 5',
 mRNA sequence.
ACCESSION BI820759
VERSION BI820759.1 GI:15932309
KEYWORDS EST.
SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 400)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11437 row: m column: 13
 High quality sequence stop: 364.

FEATURES Location/Qualifiers
 source 1. .400
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5175732"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."

BASE COUNT 73 a 131 c 131 g 65 t
 ORIGIN

Query Match 77.0%; Score 279.6; DB 12; Length 400;
 Best Local Similarity 93.9%; Pred. No. 1.5e-60;
 Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy	54	CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG	113
Db	49	CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG	108
Qy	114	GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA	173
Db	109	GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCTA	168
Qy	174	CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAAACC	233
Db	169	CATTTGGAGCCTGGACACGAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAAACC	228
Qy	234	GGGCGCGTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA	293
Db	229	GGGCGCCTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA	288
Qy	294	CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA	353

```

Db      289 CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTCCTGCAA 348
Qy      354 GTGCAGCTGA 363
          |||||
Db      349 GTGCAGCTGA 358

```

RESULT 5

BG180040

LOCUS BG180040 431 bp mRNA linear EST 06-FEB-2001

DEFINITION 602329596F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4431214 5', mRNA sequence.

ACCESSION BG180040

VERSION BG180040.1 GI:12686743

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 431)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTP

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10185 row: o column: 23

High quality sequence stop: 338.

FEATURES Location/Qualifiers

source 1..431

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4431214"

/tissue_type="adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_91"

/note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies.

Note: this is a NIH MGC Library."

BASE COUNT 95 a 126 c 148 g 62 t

ORIGIN

Query Match 77.0%; Score 279.6; DB 10; Length 431;

Best Local Similarity 93.9%; Pred. No. 1.6e-60;

Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
          || |||||

```

```

Db      28 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 87
Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
        |||||
Db      88 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCTA 147
Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAAACC 233
        |||
Db      148 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAAACC 207
Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
        |||||
Db      208 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 267
Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
        |||||
Db      268 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 327
Qy      354 GTGCAGCTGA 363
        |||||
Db      328 GTGCAGCTGA 337

```

RESULT 6

BM740537

LOCUS BM740537 461 bp mRNA linear EST 01-MAR-2002

DEFINITION K-EST0011722 S1SNU5 Homo sapiens cDNA clone S1SNU5-5-B09 5', mRNA sequence.

ACCESSION BM740537

VERSION BM740537.1 GI:19061866

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 461)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 5 row: B column: 09

High quality sequence stop: 461.

FEATURES Location/Qualifiers

source 1..461

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="S1SNU5-5-B09"

/sex="F"


```

/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-5"
/lab_host="Top10F'"
/clone_lib="S1SNU5"
/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

```

BASE COUNT 86 a 161 c 133 g 81 t
ORIGIN

Query Match 77.0%; Score 279.6; DB 12; Length 461;
Best Local Similarity 93.9%; Pred. No. 1.6e-60;
Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

Qy            54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
             || |||||
Db            148 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 207

Qy            114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
             |||||
Db            208 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCCTA 267

Qy            174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAACCC 233
             ||| |||||
Db            268 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCC 327

Qy            234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
             |||||
Db            328 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 387

Qy            294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTGCAA 353
             |||||
Db            388 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTCTCTGCAA 447

Qy            354 GTGCAGCTGA 363
             |||||
Db            448 GTGCAGCTGA 457

```

RESULT 7
BU632941/c
LOCUS BU632941 475 bp mRNA linear EST 23-SEP-2002
DEFINITION UI-H-DF0-bel-c-18-0-UI.s1 NCI_GCAP_DF0 Homo sapiens cDNA clone

UI-H-DF0-bel-c-18-0-UI 3', mRNA sequence.
 ACCESSION BU632941
 VERSION BU632941.1 GI:23300196
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 475)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 The following repetitive elements were found in this cDNA
 sequence: 97-152, >GC_rich#Low_complexity
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES Location/Qualifiers
 source 1. .475
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-DF0-bel-c-18-0-UI"
 /tissue_type="Subchondral Bone"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI_CGAP_DF0"
 /note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 NCI_CGAP_DF0 is a cDNA library containing the following
 tissue(s): Subchondral Bone. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT7T3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is GTTAAGCGTC.
 TAG_LIB=UI-H-DF0
 TAG_TISSUE=subchondral bone
 TAG_SEQ=GTTAAGCGTC"

BASE COUNT 70 a 136 c 171 g 98 t
 ORIGIN

Query Match 77.0%; Score 279.6; DB 13; Length 475;
 Best Local Similarity 93.9%; Pred. No. 1.6e-60;
 Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
 || |||||
 Db 465 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 406
 Qy 114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
 |||||
 Db 405 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCCCTA 346
 Qy 174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAAACC 233
 |||||
 Db 345 CATTTGGAGCCTGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAAACC 286
 Qy 234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
 |||||
 Db 285 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 226
 Qy 294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 353
 |||||
 Db 225 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 166
 Qy 354 GTGCAGCTGA 363
 |||||
 Db 165 GTGCAGCTGA 156

RESULT 8

AI089904/c

LOCUS AI089904 480 bp mRNA linear EST 01-OCT-1998

DEFINITION ga16a08.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:1686902 3' similar to gb:X02812_cds1 TRANSFORMING GROWTH FACTOR BETA 1 PRECURSOR (HUMAN);contains element MER22 repetitive element ;, mRNA sequence.

ACCESSION AI089904

VERSION AI089904.1 GI:3428963

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 480)

AUTHORS NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 895 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 414.

FEATURES Location/Qualifiers
 source 1. .480
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1686902"
 /tissue_type="glioblastoma (pooled)"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Brn23"
 /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT7T3 vector.
 Library is normalized, and was constructed by Bento
 Soares and M.Fatima Bonaldo."
BASE COUNT 74 a 140 c 179 g 87 t
ORIGIN

Query Match 77.0%; Score 279.6; DB 9; Length 480;
Best Local Similarity 93.9%; Pred. No. 1.6e-60;
Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||
Db      448 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 389

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      ||||| || || |||||
Db      388 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGCCCCTA 329

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAACCC 233
      ||| ||||| ||||| |||||
Db      328 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCC 269

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      ||||| ||||| ||||| |||||
Db      268 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 209

Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 353
      ||||| ||||| ||||| |||||
Db      208 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 149

Qy      354 GTGCAGCTGA 363
      |||||
Db      148 GTGCAGCTGA 139
```

RESULT 9
AW958056
LOCUS AW958056 505 bp mRNA linear EST 01-JUN-2000
DEFINITION EST370126 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.

ACCESSION AW958056
 VERSION AW958056.1 GI:8147739
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 505)
 AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
 ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
 Quackenbush,J.
 TITLE Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 JOURNAL Unpublished
 COMMENT Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 115
 Seq primer: Reverse.

FEATURES Location/Qualifiers
 source 1..505
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGE"
 /note="Vector: pBluescriptSKm"
 BASE COUNT 115 a 137 c 106 g 144 t 3 others
 ORIGIN

Query Match 77.0%; Score 279.6; DB 10; Length 505;
 Best Local Similarity 93.9%; Pred. No. 1.7e-60;
 Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 38 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 97
 Qy 114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
 ||||||||| || || ||||||||||||||||||||||||||||||||||||
 Db 98 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCCTA 157
 Qy 174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTGGCTCTGTACAACCAGCACAACCC 233
 ||| ||||||||| ||||| ||||||||||||||||||||||||||||||||
 Db 158 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCC 217
 Qy 234 GGGCGCGTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
 ||||| ||||||||||||||||||||||||||||||||||||||||||||
 Db 218 GGGCGCCTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 277
 Qy 294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 353
 ||||||||||||||||||||||||||||||||||||||||||||||||
 Db 278 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 337
 Qy 354 GTGCAGCTGA 363

Db |||||
338 GTGCAGCTGA 347

RESULT 10

BQ222205

LOCUS BQ222205 536 bp mRNA linear EST 02-MAY-2002

DEFINITION AGENCOURT_7258595 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5786680
5', mRNA sequence.

ACCESSION BQ222205

VERSION BQ222205.1 GI:20403594

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 536)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM12875 row: m column: 17

High quality sequence stop: 535.

FEATURES Location/Qualifiers

source

1. .536

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5786680"

/tissue_type="leiomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_71"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2.1 kb. "

BASE COUNT 113 a 193 c 150 g 77 t 3 others

ORIGIN

Query Match 77.0%; Score 279.6; DB 13; Length 536;

Best Local Similarity 93.9%; Pred. No. 1.7e-60;

Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113

|| |||||

Db 65 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 124

Qy 114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173

||||||| || || |||||

Db 125 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCTA 184

[illegible]

RESULT 11

BM670402/c

LOCUS BM670402 538 bp mRNA linear EST 27-FEB-2002

DEFINITION UI-E-DW1-ahc-1-03-0-UI.s1 UI-E-DW1 Homo sapiens cDNA clone

UI-E-DW1-ahc-1-03-0-UI 3', mRNA sequence.

ACCESSION BM670402

VERSION BM670402.1 GI:18980299

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 538)

AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa

DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

The following repetitive elements were found in this cDNA

sequence: 97-152, >GC rich#Low complexity

Seq primer: M13 Forward

POLYA=Yes .

FEATURES	Location/Qualifiers
----------	---------------------

```
source      1.  .538
```

```
/organism="Homo sapiens"
```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DW1-ahc-1-03-0-UI"
/tissue_type="lens"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DW1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-DW1 is a normalized cDNA library containing the
following tissue(s): lens. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGATTAGCGA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG_LIB=UI-E-DW1
TAG_TISSUE=human lens
TAG_SEQ=CGATTAGCGA"

```

```

BASE COUNT      78 a      153 c      195 g      112 t
ORIGIN

```

```

Query Match          77.0%;  Score 279.6;  DB 12;  Length 538;
Best Local Similarity 93.9%;  Pred. No. 1.7e-60;
Matches 291;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      465 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 406

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      ||||||||| || || ||||||||||||||||||||||||||||||||||||
Db      405 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCTA 346

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTCTGGCTCTGTACAACCAGCACAAACC 233
      ||| |||||||| ||||| ||||||| ||||||| ||||||| ||||||| |||||
Db      345 CATTTGGAGCCTGGACACGAGTACAGCAAGGTCTTGGCCCTGTACAACCAGCATAAACC 286

Qy      234 GGGCGCGTCGGCGGCGCGGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      285 GGGCGCCTCGGCGGCGCGGTGCTGCGTGCCGCGAGGCGCTGGAGCCGCTGCCCATCGTGTA 226

Qy      294 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 353
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      225 CTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGCAA 166

Qy      354 GTGCAGCTGA 363
      |||||||||
Db      165 GTGCAGCTGA 156

```


RESULT 12
 BI834732
 LOCUS BI834732 540 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603090326F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5229246 5',
 mRNA sequence.
 ACCESSION BI834732
 VERSION BI834732.1 GI:15946269
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 540)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11577 row: c column: 07
 High quality sequence stop: 535.
 FEATURES
 source Location/Qualifiers
 1..540
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5229246"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_120"
 /note="Organ: pooled pancreas and spleen; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of spleen and pancreas from 28 yo
 male. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.5 kb, insert size range 1-2.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 025. Note: this is a NIH_MGC Library."
 BASE COUNT 104 a 188 c 154 g 94 t
 ORIGIN

Query Match 77.0%; Score 279.6; DB 12; Length 540;
 Best Local Similarity 93.9%; Pred. No. 1.7e-60;
 Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
 || |||||
 Db 220 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 279
 Qy 114 GAAGTGGATTTCATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
 ||||| || || |||||

```

Db      280 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTA 339
Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCCTGGCTCTGTACAACCAGCACAAACC 233
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      340 CATTTGGAGCCTGGACACGCGAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAAACC 399
Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      400 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCGAGGCGCTGGAGCCGCTGCCCATCGTGTA 459
Qy      294 CTACGTGGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAA 353
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      460 CTACGTGGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCTTGCAA 519
Qy      354 GTGCAGCTGA 363
      ||||| |||||
Db      520 GTGCAGCTGA 529

```

RESULT 13

BQ016027/c

LOCUS BQ016027 587 bp mRNA linear EST 17-JUN-2002

DEFINITION UI-H-DT1-avz-h-20-0-UI.s1 NCI_CGAP_DT1 Homo sapiens cDNA clone
IMAGE:5886403 3', mRNA sequence.

ACCESSION BQ016027

VERSION BQ016027.1 GI:19751304

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 587)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

The following repetitive elements were found in this cDNA

sequence: 97-152, >GC-rich#Low_complexity

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES Location/Qualifiers

source 1. .587

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5886403"

/tissue_type="Metastatic Chondrosarcoma"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

```

/clone_lib="NCI_CGAP_DT1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI_CGAP_DT1 is a normalized cDNA library containing the
following tissue(s): Metatastic Chondrosarcoma in Lung.
The library was constructed according to Bonaldo, Lennon
and Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is AACTGTTCGG.
TAG_LIB=UI-H-DT1
TAG_TISSUE=lung metastatic chondrosarcoma
TAG_SEQ=AACTGTTCGG"

```

```

BASE COUNT      91 a      163 c      213 g      120 t
ORIGIN

```

```

Query Match          77.0%;  Score 279.6;  DB 12;  Length 587;
Best Local Similarity 93.9%;  Pred. No. 1.8e-60;
Matches 291;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      465 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 406

Qy      114 GAAGTGGATTTCATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      ||||||||| || || |||||||||||||||||||||||||||||||||||
Db      405 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCCTA 346

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTTGGCTCTGTACAACCAGCACAACCC 233
      ||| ||||||||| ||||| ||||||| ||||||| ||||||| |||||||
Db      345 CATTTGGAGCCTGGACACGCGAGTACAGCAAGGTCTTGGCCCTGTACAACCAGCATAACCC 286

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCGGCGCTGGAGCCACTGCCCATCGTGTA 293
      ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      285 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCGGCGCTGGAGCCGCTGCCCATCGTGTA 226

Qy      294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTGCAA 353
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db      225 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTCTCTGCAA 166

Qy      354 GTGCAGCTGA 363
      |||||||||
Db      165 GTGCAGCTGA 156

```

```

RESULT 14
BQ003136/c
LOCUS      BQ003136              593 bp      mRNA      linear      EST 26-MAR-2002
DEFINITION UI-H-EI1-ayx-d-19-0-UI.s1 NCI_CGAP_EI1 Homo sapiens cDNA clone
            IMAGE:5844834 3', mRNA sequence.
ACCESSION  BQ003136
VERSION    BQ003136.1  GI:19728036

```

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 593)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 The following repetitive elements were found in this cDNA
 sequence: 97-152, >GC_rich#Low_complexity
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES Location/Qualifiers
 source 1. .593
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5844834"
 /tissue_type="Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI_CGAP_EI1"
 /note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
 with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
 NCI_CGAP_EI1 is a normalized cDNA library containing the
 following tissue(s): Chondrosarcoma. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT7T3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 ACACTTGACAC.
 TAG_LIB=UI-H-EI1
 TAG_TISSUE=chondrosarcoma
 TAG_SEQ=ACACTTGACAC"

BASE COUNT 91 a 164 c 217 g 121 t
 ORIGIN

Query Match 77.0%; Score 279.6; DB 12; Length 593;
 Best Local Similarity 93.9%; Pred. No. 1.8e-60;
 Matches 291; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113

```

      || |||||
Db      465 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 406
QY      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      |||||
Db      405 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCCCTA 346
QY      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTTGGCTCTGTACAACCAGCACAAACC 233
      |||||
Db      345 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGGCCCTGTACAACCAGCATAAACC 286
QY      234 GGGCGCGTCGGCGGCGCGCTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      |||||
Db      285 GGGCGCCTCGGCGGCGCGCTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 226
QY      294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 353
      |||||
Db      225 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAA 166
QY      354 GTGCAGCTGA 363
      |||||
Db      165 GTGCAGCTGA 156

```

RESULT 15

AI951831/c

LOCUS AI951831 598 bp mRNA linear EST 06-SEP-1999

DEFINITION wx38b08.x1 NCI_CGAP_Pit1 Homo sapiens cDNA clone IMAGE:2545911 3' similar to gb:X02812_cds1 TRANSFORMING GROWTH FACTOR BETA 1 PRECURSOR (HUMAN);contains PTR7.t3 MER22 repetitive element ;, mRNA sequence.

ACCESSION AI951831

VERSION AI951831.1 GI:5744141

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 598)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco

High quality sequence stop: 425.

FEATURES Location/Qualifiers

source 1. .598

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2545911"
/tissue_type="four pooled pituitary adenomas"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pit1"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies."

```

```

BASE COUNT      95 a      173 c      222 g      107 t      1 others
ORIGIN

```

```

Query Match          77.0%;  Score 279.6;  DB 9;  Length 598;
Best Local Similarity 93.9%;  Pred. No. 1.8e-60;
Matches 291;  Conservative 0;  Mismatches 19;  Indels 0;  Gaps 0;

```

```

Qy      54 CAAGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTG 113
      || |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      447 CACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTG 388

Qy      114 GAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTA 173
      ||||||||| || || |||||||||||||||||||||||||||||||||||||
Db      387 GAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCTGCCCCTA 328

Qy      174 CATCTGGAGCCTAGACACTCAGTACAACAAGGTCTCTGGCTCTGTACAACCAGCACAACCC 233
      ||| ||||||||| ||||| |||||||||||||||||||||||||||||||||
Db      327 CATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCC 268

Qy      234 GGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTA 293
      ||||||| |||||||||||||||||||||||||||||||||||||||||
Db      267 GGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTA 208

Qy      294 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 353
      |||||||||||||||||||||||||||||||||||||||||||||||||
Db      207 CTACGTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAA 148

Qy      354 GTGCAGCTGA 363
      |||||||||
Db      147 GTGCAGCTGA 138

```

```

Search completed: October 28, 2003, 00:08:15
Job time : 902.804 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2003, 19:11:58 ; Search time 10.0108 Seconds
(without alignments)
1902.657 Million cell updates/sec

Title: US-10-017-372E-15
Perfect score: 685
Sequence: 1 ALDTNYCFSSTDYKDDDDKE.....GRKPKVEQLSNMIVRSCKCS 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	624	91.1	112	11	AAR04076		Sequence of mature
2	624	91.1	112	11	AAR08142		Platelet-derived h
3	624	91.1	112	14	AAR43263		TGF-beta. Homo sa
4	624	91.1	112	14	AAR42311		Recombinant human
5	624	91.1	112	17	AAR92773		Human TGF-beta 1.
6	624	91.1	112	17	AAR91956		Human transforming
7	624	91.1	112	18	AAW08173		TGF-beta1 active f
8	624	91.1	112	19	AAW78781		Human transforming
9	624	91.1	112	20	AAW08299		Human growth facto
10	624	91.1	112	20	AAW97091		The mature form of
11	624	91.1	112	20	AAW84207		Transforming growt
12	624	91.1	112	21	AAW92010		Human transforming
13	624	91.1	112	21	AAW67950		Human transforming
14	624	91.1	112	22	AAB35937		TGF-beta 1 amino a
15	624	91.1	112	23	AAM51939		Human TGFbeta prot
16	624	91.1	112	24	ABU08656		Human transforming
17	624	91.1	115	22	AAB73204		TGF-beta1 C-termin
18	624	91.1	115	24	ABG76033		Human TGF-beta 1 C
19	624	91.1	118	18	AAW08176		TGF-beta fusion pr
20	624	91.1	122	18	AAW08181		TGF-beta fusion pr
21	624	91.1	124	18	AAW08177		TGF-beta fusion pr
22	624	91.1	128	18	AAW08188		TGF-beta fusion pr
23	624	91.1	130	18	AAW08185		TGF-beta fusion pr
24	624	91.1	131	18	AAW08182		TGF-beta fusion pr
25	624	91.1	132	18	AAW08189		TGF-beta fusion pr
26	624	91.1	133	18	AAW08178		TGF-beta fusion pr
27	624	91.1	134	18	AAW08187		TGF-beta fusion pr
28	624	91.1	134	18	AAW08190		TGF-beta fusion pr
29	624	91.1	135	18	AAW08183		TGF-beta fusion pr
30	624	91.1	136	18	AAW08186		TGF-beta fusion pr
31	624	91.1	137	18	AAW08180		TGF-beta fusion pr
32	624	91.1	137	18	AAW08184		TGF-beta fusion pr
33	624	91.1	139	18	AAW08179		TGF-beta fusion pr
34	624	91.1	390	7	AAP61468		PreTGF-beta gene p
35	624	91.1	390	11	AAR04034		Sequence of pre-TG
36	624	91.1	390	11	AAR05258		Human pre-transfor
37	624	91.1	390	12	AAR13813		Human pro-TGF-beta
38	624	91.1	390	13	AAR20124		Sequence of simian
39	624	91.1	390	15	AAR46227		Human pre-TGF-beta
40	624	91.1	390	16	AAR73596		Human TGF-beta 1 p
41	624	91.1	390	17	AAR90827		Pre-transforming g
42	624	91.1	390	19	AAW78785		Human pre-transfor
43	624	91.1	390	22	AAM39186		Human polypeptide
44	624	91.1	390	23	AAU77101		Human transforming
45	624	91.1	390	23	AAE13596		Porcine transformi

ALIGNMENTS

RESULT 1

AAR04076

ID AAR04076 standard; protein; 112 AA.

XX

AC AAR04076;

XX

DT 25-MAR-2003 (updated)

DT 31-OCT-2002 (updated)

DT 31-MAY-1989 (first entry)

XX

DE Sequence of mature human TGF-beta 1.

XX

KW Transforming growth factor beta-3 (TGF beta 3); tumour cells; growth inhibition.

XX

OS Homo sapiens.

XX

PN WO8912101-A.

XX

PD 14-DEC-1989.

XX

PF 08-JUN-1988; 88WO-US01945.

XX

PR 08-JUN-1988; 88WO-US01945.

XX

PA (GETH) GENENTECH INC.

XX

PI Dernyck RMA, Goeddel DV;

XX

DR WPI; 1990-007474/01.

XX

PT Nucleotide sequence encoding transforming growth factor beta-3 used as a probe, or to produce TGF beta 3, for inhibiting growth of certain normal and neoplastic cells, eg A549.

XX

PS Disclosure; Fig. 3; 6lpp; English.

XX

CC This sequence is that of mature human transforming growth factor-beta 1 (TGF-beta 1) polypeptide. It differs from human and porcine TGF-beta subtype 3 sequences at positions 9-11,13,19,40,45,52,57,58,60,63,67,68, 71,75,82,87,and 95. In addition porcine TGF-beta 3 has an AA substitution at posn. 96. The N-terminal AA sequence of bovine TGF-beta 2 differs from that of human TGF-beta 1 at posns.4,5,9-13,and 19.

CC The nucleic acid sequence encoding the second subtype of TGF-beta (TGF-beta 3)is useful as a probe or to produce TGF-beta 3 for both normal and neoplastic cell growth inhibition.

CC (Updated on 31-OCT-2002 to add missing OS field.)

CC (Updated on 25-MAR-2003 to correct PR field.)

CC (Updated on 25-MAR-2003 to correct PI field.)

XX

SQ Sequence 112 AA;

Query Match 91.1%; Score 624; DB 11; Length 112;

Best Local Similarity 93.3%; Pred. No. 1.9e-59;

Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60

```

Db          |||||
1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52

Qy          61 SLDTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
          |||||

Db          53 SLDTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 112

```

RESULT 2

AAR08142

ID AAR08142 standard; protein; 112 AA.

XX

AC AAR08142;

XX

DT 25-MAR-2003 (updated)

DT 28-FEB-1991 (first entry)

XX

DE Platelet-derived human TGF-beta monomer.

XX

KW Transforming growth factor; cartilages inducing factor; CIF;

KW polycythemia; thrombocytosis; splenomegaly.

XX

OS Homo sapiens.

XX

PN US4971952-A.

XX

PD 20-NOV-1990.

XX

PF 27-OCT-1988; 88US-0263635.

XX

PR 27-OCT-1988; 88US-0263635.

XX

PA (CLGE) COLLAGEN CORP.

XX

PI Bentze H, Ellingsworth L, Armstrong R;

XX

DR WPI; 1990-368150/49.

XX

PT Treating inflammation associated with transplants, etc. - using

PT cartilage inducing factor polypeptide homodimer cpd.

XX

PS Disclosure; Fig 1; 14pp; English.

XX

CC The TGF-beta may be used to derive cartilage inducing factors (CIFs)

CC useful for inhibition of acute/chronic inflammation, treatment of red

CC blood cell dysfunction associated with polycythemia, thrombocytosis

CC or splenomegaly.

CC (Updated on 25-MAR-2003 to correct PR field.)

XX

SQ Sequence 112 AA;

Query Match 91.1%; Score 624; DB 11; Length 112;

Best Local Similarity 93.3%; Pred. No. 1.9e-59;

Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```

Qy          1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
          |||||

```

Db 1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52

Qy 61 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120

Db 53 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

RESULT 3

AAR43263

ID AAR43263 standard; protein; 112 AA.

XX

AC AAR43263;

XX

DT 25-MAR-2003 (updated)

DT 27-APR-1994 (first entry)

XX

DE TGF-beta.

XX

KW TGF; transforming growth factor; scar formation; CNS;

KW central nervous system; infusion; antagonist; brain;

KW spinal cord; regeneration; decorin; Arg-Gly-Asp.

XX

OS Homo sapiens.

XX

PN WO9319783-A1.

XX

PD 14-OCT-1993.

XX

PF 01-APR-1993; 93WO-US03068.

XX

PR 01-APR-1992; 92US-0860704.

XX

PA (WHIT-) WHITTIER INST DIABETES & ENDOCRINOLOGY.

XX

PI Baird A, Logan A;

XX

DR WPI; 1993-336598/42.

XX

PT Inhibiting or enhancing scar formation in the CNS - by infusion

PT of transforming growth factor-beta antagonists, e.g.

PT anti-(TGF-beta)-antibodies, or TGF-beta itself, respectively

XX

PS Disclosure; Page 5; 32pp; English.

XX

CC The mature form of TGF-beta is comprised of two identical chains,
 CC each of 112 amino acids. Since TGF-beta is a regulator of scar
 CC prodn. after a penetrating injury to the brain or spinal cord, which
 CC scar prodn. can preclude neuronal recovery, TGF-beta antagonists
 CC have potential use as adjuncts to those therapies designed to
 CC promote regeneration and reconnection of damaged neural pathways.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 112 AA;

Query Match 91.1%; Score 624; DB 14; Length 112;

Best Local Similarity 93.3%; Pred. No. 1.9e-59;

Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```

QY      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      |||||
Db      1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52

QY     61 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      |||||
Db     53 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 112

```

RESULT 4

AAR42311

ID AAR42311 standard; peptide; 112 AA.

XX

AC AAR42311;

XX

DT 01-APR-1994 (first entry)

XX

DE Recombinant human TGF-beta.

XX

KW Transforming growth factor beta; stable; treatment; wounds;

KW fractures; inflammation; autoimmune disease; tumours.

XX

OS Homo sapiens (recombinant).

XX

PN JP05252956-A.

XX

PD 05-OCT-1993.

XX

PF 11-JUL-1991; 91JP-0171280.

XX

PR 11-JUL-1991; 91JP-0171280.

XX

PA (EART) EARTH SEIYAKU KK.

XX

DR WPI; 1993-347572/44.

XX

PT Prodn. of recombinant cell for high expression of polypeptide

PT such as human TGF-beta - for treatment of e.g. wounds, fractures,

PT inflammation, auto-immune diseases or tumours

XX

PS Disclosure; Page 15; 18pp; Japanese.

XX

CC Human transforming growth factor (TGF) -beta can be produced in

CC recombinant cells, and thus produced is stable in the presence of 50

CC microg/ml methotrexate. TGF-beta may be produced in large amounts

CC via recombinant means and has uses such a wound healing, treatment

CC of fractures, inflammation, autoimmune disease or tumours.

XX

SQ Sequence 112 AA;

Query Match 91.1%; Score 624; DB 14; Length 112;

Best Local Similarity 93.3%; Pred. No. 1.9e-59;

Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```

QY      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      |||||

```

```

Db          1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52
Qy          61 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
              |||
Db          53 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

```

RESULT 5

AAR92773

ID AAR92773 standard; Protein; 112 AA.

XX

AC AAR92773;

XX

DT 17-JUL-1996 (first entry)

XX

DE Human TGF-beta 1.

XX

KW Transforming growth factor type beta; TGF-beta 1;

KW protein renaturation; protein folding.

XX

OS Homo sapiens.

XX

PN WO9603433-A1.

XX

PD 08-FEB-1996.

XX

PF 12-JUL-1995; 95WO-EP02719.

XX

PR 25-JUL-1994; 94EP-0810439.

XX

PA (CIBA) CIBA GEIGY AG.

XX

PI Cerletti N;

XX

DR WPI; 1996-117000/12.

DR N-PSDB; AAT17235.

XX

PT Prodn. of dimeric biologically active transforming growth factor -

PT by refolding denatured monomer in detergent-free folding buffer

PT contg. specific organic solvent to improve yield

XX

PS Example 1B; Page 29-30; 54pp; English.

XX

CC Non-soluble, monomeric transforming growth factor TGF-beta 1

CC (AAR92773) was recovered from E. coli LC 137 (DSM 5658) transformants

CC carrying plasmid pPLMu.TGF-beta 1, which includes the coding

CC sequence (AAT17235) for TGF-beta 1. A biologically active, dimeric

CC form of TGF-beta 1 was obtd. by refolding this monomer in detergent-

CC free buffer contg. DMSO and/or DMF. Dimers of TGF-beta 2 (AAR92774)

CC and TGF-beta 3 (AAR92772), and hybrid dimers (see also AAR92775-77),

CC were also produced.

XX

SQ Sequence 112 AA;

Query Match 91.1%; Score 624; DB 17; Length 112;

Best Local Similarity 93.3%; Pred. No. 1.9e-59;

Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52

Qy 61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
 |||||
 Db 53 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 112

RESULT 6

AAR91956

ID AAR91956 standard; Protein; 112 AA.

XX

AC AAR91956;

XX

DT 10-JUN-1996 (first entry)

XX

DE Human transforming growth factor beta 1.

XX

KW Transforming growth factor beta; TGF; regulator; method;
 KW proliferation; differentiation; wound healing; solvent.

XX

OS Homo sapiens.

XX

PN WO9603432-A1.

XX

PD 08-FEB-1996.

XX

PF 12-JUL-1995; 95WO-EP02718.

XX

PR 25-JUL-1994; 94EP-0810438.

XX

PA (CIBA) CIBA GEIGY AG.

XX

PI Cerletti N;

XX

DR WPI; 1996-116999/12.

DR N-PSDB; AAT15462.

XX

PT Prodn. of dimeric, biologically active transforming growth factor
 PT beta - by refolding denatured monomer in buffer contg. mild
 PT detergent and specific organic solvents to improve yields

XX

PS Example 1; Page 33; 59pp; English.

XX

CC AAR91956-R91958 are the amino acid sequences of human transforming
 CC growth factor (TGF) beta-1, TGF beta-2 and TGF beta-3 which are used to
 CC produce TGF beta-like proteins in dimeric form. The TGF beta-like
 CC proteins produced are hybrids of 2 different types of TGF beta e.g.
 CC TGF beta-1-3, TGF beta-2-3, etc, or bone morphogenic proteins e.g.
 CC BMP-2. The TGF beta hybrids were made using a new process of producing
 CC dimeric, biologically active TGF beta-like proteins. The new process
 CC involves treating denatured TGF beta monomers with folding buffer contg.
 CC a mild detergent (CHAPS, CHAPSO or digitonin) and at least one of the
 CC solvents DMSO (dimethyl sulphoxide), DMSO2 (dimethylsulphone) and DMF
 CC (dimethyl formamide). The detergent allows folding of the monomer such

XX
 PS Claim 12; Page 45-46; 59pp; English.
 XX
 CC A novel transforming growth factor-beta (TGF-beta) fusion protein
 CC comprises a purification tag and a TGF active fragment. The present
 CC sequence represents a specifically claimed TGF-beta1 active fragment.
 CC Additionally, the fusion protein may comprise proteinase-sensitive
 CC linker sites and binding domain so the protein sequence may contain
 CC some or all of the following elements: purification tag:proteinase
 CC site:ECM binding site:proteinase site:TGF-beta. TGF-beta promotes
 CC wound healing, and the fusion protein can be used to reduce surgery
 CC recovery time and in the preparation of artificial skin. The inclusion
 CC of a purification tag facilitates purification of the fusion protein.
 CC The proteinase site is included to permit cleavage and release of the
 CC purification tag after purification if desired. The extracellular
 CC matrix binding site facilitates delivery of the fusion protein to the
 CC desired site of action. Delivery of the TGF-beta to the site to be
 CC treated reduces the amount of TGF-beta required to be administered to
 CC be effective and reduces the concentration of circulating TGF-beta
 CC which may result in undesirable effects.
 XX
 SQ Sequence 112 AA;

Query Match 91.1%; Score 624; DB 18; Length 112;
 Best Local Similarity 93.3%; Pred. No. 1.9e-59;
 Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52
 QY 61 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
 |||||
 Db 53 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

RESULT 8

AAW78781

ID AAW78781 standard; Protein; 112 AA.

XX

AC AAW78781;

XX

DT 25-MAR-2003 (updated)

DT 21-DEC-1998 (first entry)

XX

DE Human transforming growth factor-beta 1 N-terminal sequence.

XX

KW Transforming growth factor-beta 1; TGF-beta 1; human.

XX

OS Homo sapiens.

XX

PN US5801231-A.

XX

PD 01-SEP-1998.

XX

PF 30-MAY-1995; 95US-0454468.

XX

KW S-sulphonated TGF-beta-like protein; wound treatment; cancer;
 KW bone repair; tissue repair; bone marrow protective agent;
 KW cardioprotection; anti-inflammatory; immunosuppressive;
 KW ulcer; bed sore.
 XX
 OS Homo sapiens.
 XX
 PN EP891985-A1.
 XX
 PD 20-JAN-1999.
 XX
 PF 27-NOV-1990; 98EP-0113487.
 XX
 PR 06-DEC-1989; 89GB-0027546.
 PR 27-NOV-1990; 90EP-0810922.
 XX
 PA (NOVS) NOVARTIS AG.
 XX
 PI Cerletti N, Cox D, McMaster GK, Meyhack B, Schmitz A;
 XX
 DR WPI; 1999-083520/08.
 DR N-PSDB; AAX15245.
 XX
 PT Producing biologically active dimeric Transforming Growth
 PT Factor-beta - by refolding new monomeric Transforming Growth
 PT Factor-beta, useful for treatment of wounds and cancer
 XX
 PS Example 1; Page 28; 32pp; English.
 XX
 CC The present sequence represents the mature form of transforming growth
 CC factor-beta-1. Dimeric, biologically active TGF-beta-like protein
 CC can be produced by subjecting the denatured monomeric form to refolding
 CC conditions. The new monomeric S-sulphonated TGF-beta-like protein is
 CC useful for the production of the dimeric, biologically active
 CC TGF-beta-like protein, which is useful for the treatment of wounds
 CC (surface or internal) and cancer in a mammal, in bone and tissue
 CC repair, as a bone marrow protective agent, a mediator of
 CC cardioprotection, for the production of an anti-inflammatory or
 CC immunosuppressive preparation. Treatment is useful for animals,
 CC especially humans, and wound treatment (e.g. ulcers, bed sores etc.) is
 CC particularly useful for the elderly.
 CC (Updated on 20-MAR-2003 to correct PF field.)
 CC (Updated on 20-MAR-2003 to correct PR field.)
 XX
 SQ Sequence 112 AA;

Query Match 91.1%; Score 624; DB 20; Length 112;
 Best Local Similarity 93.3%; Pred. No. 1.9e-59;
 Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52
 Qy 61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
 |||||
 Db 53 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

RESULT 11

AAW84207

ID AAW84207 standard; Protein; 112 AA.

XX

AC AAW84207;

XX

DT 25-MAR-1999 (first entry)

XX

DE Transforming growth factor beta active fragment.

XX

KW Proteinase site; bone morphogenetic fusion protein; bone binding site;

KW bone morphogenetic protein; transforming growth factor beta;

KW active fragment; wound healing; bone growth; purification tag.

XX

OS Homo sapiens.

XX

PN WO9855137-A1.

XX

PD 10-DEC-1998.

XX

PF 02-JUN-1998; 98WO-US11189.

XX

PR 03-JUN-1997; 97US-0868452.

XX

PA (HALL/) HALL F L.

PA (HANB/) HAN B.

PA (NIMN/) NIMNI M E.

PA (SHOR/) SHORS E C.

PA (WULL/) WU L.

XX

PI Hall FL, Han B, Nimni ME, Shors EC, Wu L;

XX

DR WPI; 1999-059875/05.

DR N-PSDB; AAV99375.

XX

PT New bone morphogenetic fusion proteins - comprising a purification

PT tag and a bone morphogenetic active fragment, used for enhancing

PT wound healing or bone growth

XX

PS Example 1; Page 41; 64pp; English.

XX

CC The present sequence represents a transforming growth factor beta active

CC fragment. The protein can be used in place of a bone morphogenetic

CC active fragment to create the fusion proteins of the invention. When a

CC bone morphogenetic active fragment is used, the fusion proteins are

CC designated bone morphogenetic fusion proteins. The bone morphogenetic

CC fusion protein may contain some or all of the following elements: a

CC purification tag, a proteinase site, an ECM/bone binding site, a second

CC proteinase site, and a bone morphogenetic protein active fragment.

CC The bone morphogenetic fusion proteins can be used for enhancing wound

CC healing or bone growth.

XX

SQ Sequence 112 AA;

Query Match

91.1%; Score 624; DB 20; Length 112;

Best Local Similarity 93.3%; Pred. No. 1.9e-59;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
        |||||
Db      1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52

Qy     61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
        |||||
Db     53 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112
```

RESULT 12

AA92010

ID AAY92010 standard; Protein; 112 AA.

XX

AC AAY92010;

XX

DT 19-JUL-2000 (first entry)

XX

DE Human transforming growth factor beta 1 monomer.

XX

KW human transforming growth factor beta 1 monomer; CKGF; mutant;

KW cystine knot growth factor; hairpin loop; infertility.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 1..20

FT /note= "optionally mutated to increase electrostatic
FT interaction between beta hairpin structure and
FT a receptor"

FT Domain 21..40

FT /label= beta_hairpin_loop_1

FT /note= "mutant optionally comprises one or more
FT substitutions in these residues"

FT Misc-difference 41..81

FT /note= "optionally mutated to increase electrostatic
FT interaction between beta hairpin structure and
FT a receptor"

FT Domain 82..102

FT /label= beta_hairpin_loop_3

FT /note= "mutant optionally comprises one or more
FT substitutions in these residues"

FT Misc-difference 103..112

FT /note= "optionally mutated to increase electrostatic
FT interaction between beta hairpin structure and
FT a receptor"

XX

PN WO200017360-A1.

XX

PD 30-MAR-2000.

XX

PF 19-MAR-1999; 99WO-US05908.

XX

PR 22-SEP-1998; 98WO-US19772.

XX

PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Weintraub BD, Szkudlinski MW;
 XX
 DR WPI; 2000-283585/24.
 XX
 PT New mutant cystine knot growth factor proteins comprising one or more
 PT mutant subunits, useful for treating or preventing diseases e.g.
 PT hypothyroidism and thyroid cancer
 XX
 PS Claim 193; Page 300; 320pp; English.
 XX
 CC This is the wild type human transforming growth factor beta 1 monomer.
 CC Mutants comprise at least one electrostatic charge altering mutation in a
 CC beta hairpin loop, resulting in increased bioactivity.
 CC Mutant cystine knot growth factor (CKGF) proteins comprising one or more
 CC mutant subunits and having novel properties or improved pharmacological
 CC properties, compared to wild type CKGFs, are claimed. The CKGF
 CC superfamily comprises at least four families of growth factors: the
 CC glycoprotein hormones, the platelet-derived growth factor (PDGF) family,
 CC the neurotrophins and the transforming growth factor-beta family; the
 CC families are known to be structurally similar (especially comprising the
 CC cystine knot topology) and it was shown that mutations at certain
 CC positions in the CKGF hairpin loops of family members and other members
 CC of the CKGF superfamily could significantly alter the biological
 CC activities of the CKGF.
 CC Mutant transforming growth factor family proteins or analogues are useful
 CC for treatment of ovulatory dysfunction, luteal phase defect, unexplained
 CC infertility, time-limited conception and in assisted reproduction.
 XX
 SQ Sequence 112 AA;

Query Match 91.1%; Score 624; DB 21; Length 112;
 Best Local Similarity 93.3%; Pred. No. 1.9e-59;
 Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52
 QY 61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
 |||||
 Db 53 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

RESULT 13

AA67950

ID AAY67950 standard; protein; 112 AA.

XX

AC AAY67950;

XX

DT 03-APR-2000 (first entry)

XX

DE Human transforming growth factor SEQ ID NO:2.

XX

KW Human; transforming growth factor; Pax4 gene; expression; potentiator;

KW insulin; pancreatic beta cell; diabetes.

XX
 OS Homo sapiens.
 XX
 PN WO9966073-A1.
 XX
 PD 23-DEC-1999.
 XX
 PF 15-JUN-1999; 99WO-JP03182.
 XX
 PR 16-JUN-1998; 98JP-0167976.
 XX
 PA (YAMA) YAMANOUCHI PHARM CO LTD.
 XX
 PI Ueda Y;
 XX
 DR WPI; 2000-097752/08.
 XX
 PT Screening potential Pax4 gene potentiators, used in treatment of, e.g.
 PT diabetes -
 XX
 PS Disclosure; Page 29; 38pp; Japanese.
 XX
 CC The present invention describes the a method for screening potential
 CC inhibitors of the expression of the Pax4 gene by contacting the
 CC potential inhibitor with pancreatic beta cells and measuring the
 CC expression of the gene in these cells is new. Substances identified
 CC by the screening method potentiate the expression of the Pax4 gene in
 CC pancreatic beta cells and accelerate the expression of insulin gene in
 CC those cells. The method can be used in the treatment of disorders in
 CC which the exhaustion of pancreatic beta cells is involved, such as
 CC diabetes. The present sequence represents the human transforming growth
 CC factor protein which is used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 112 AA;

Query Match 91.1%; Score 624; DB 21; Length 112;
 Best Local Similarity 93.3%; Pred. No. 1.9e-59;
 Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy	1	ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW	60
Db	1	ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW	52
Qy	61	SLDTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	120
Db	53	SLDTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	112

RESULT 14

AAB35937

ID AAB35937 standard; protein; 112 AA.

XX

AC AAB35937;

XX

DT 26-FEB-2001 (first entry)

XX

DE TGF-beta 1 amino acid sequence.
XX
KW Heparin binding; vascular graft; matrix; cell adhesion; growth factor;
KW wound healing; dermal wound; wound healing; TGF-beta 1.
XX
OS Unidentified.
XX
PN WO200064481-A1.
XX
PD 02-NOV-2000.
XX
PF 22-APR-1999; 99WO-IB00800.
XX
PR 22-APR-1999; 99WO-IB00800.
XX
PA (ETHZ-) ETH ZURICH & UNIV ZURICH.
XX
PI Sakiyama SE, Hubbell JA;
XX
DR WPI; 2001-024627/03.
XX
PT Matrix for controlled release of growth factor for wound healing, has
PT substrate that attaches heparin binding peptide, protein growth factor
PT that bind heparin with low affinity, and heparin or heparin-like
PT polymer -
XX
PS Example 5; Page 21; 48pp; English.
XX
CC This invention relates to a matrix comprising a substrate capable of
CC providing attachment of a heparin binding peptide (HBP), a peptide
CC comprising a binding domain which binds heparin with high affinity,
CC heparin or heparin-like polymer, and a protein growth factor or peptide
CC fragment which has a domain that binds heparin with low affinity.
CC Included in the invention is a vascular graft comprising the matrix,
CC which is capable of supporting cell adhesion. The matrix is used for
CC delivering low heparin binding affinity growth factor proteins or
CC peptides in a controlled manner suitable for wound healing. The matrix
CC can be used in an article for treating dermal wounds, and in an
CC implantable sterilized composition capable of supporting cell adhesion.
CC The present sequence represents a growth factor protein. The protein is
CC used in an example illustrating that non-heparin-binding growth factors
CC can be released in a controlled manner from heparin-based drug delivery
CC systems based on their low affinity for heparin.
XX
SQ Sequence 112 AA;

AAM51939

XX

XX

XX

XX

KW agonist: ectopic bone formation: psoriasis: muscular atrophy: scar:

KW antifibrotic: hepatotropic: vulnerary: TGFbeta1.

OS

PN

PD

PF

PR

PA

PI

DR

PT

PT for receptor binding -

PS

CC

CC activity, and where the mutation results in the protein binding with low

CC growth factor beta (TGFbeta) superfamily. The mutant sequences of the

CC overexpression of TGFbeta family proteins, including ectopic bone

CC cirrhosis. The present sequence is the human TGFbeta1 protein.

SQ

Query Match

Matches 112; Conservative 0; Mismatches 0; Indels

QY 1 ALDINYCFSSFDYRDDDDRENRCCVRQLYIDFRKDLGW
| | | | | | | | | | | | | | | | | | | | |

```

Db          1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52
Qy          61 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          53 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

```

Search completed: October 28, 2003, 09:06:47
Job time : 10.0108 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 09:09:54 ; Search time 7.28061 Seconds
(without alignments)
2760.110 Million cell updates/sec

Title: US-10-017-372E-15
Perfect score: 685
Sequence: 1 ALDTNYCFSSTDYKDDDDKE.....GRKPKVEQLSNMIVRSCKCS 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	624	91.1	112	10	US-09-813-271B-2	Sequence 2, Appli
2	624	91.1	113	10	US-09-813-398-13	Sequence 13, Appl
3	624	91.1	114	10	US-09-813-459-22	Sequence 22, Appl
4	624	91.1	114	14	US-10-115-406-21	Sequence 21, Appl
5	624	91.1	114	15	US-10-154-333-23	Sequence 23, Appl
6	624	91.1	115	10	US-09-859-211-47	Sequence 47, Appl
7	624	91.1	115	10	US-09-880-708-25	Sequence 25, Appl
8	624	91.1	115	11	US-09-872-856-47	Sequence 47, Appl
9	624	91.1	115	15	US-10-335-483-29	Sequence 29, Appl
10	624	91.1	139	14	US-10-002-278-8	Sequence 8, Appli
11	624	91.1	315	11	US-09-214-592-25	Sequence 25, Appl
12	624	91.1	390	10	US-09-756-283A-23	Sequence 23, Appl
13	624	91.1	390	11	US-09-214-592-26	Sequence 26, Appl
14	624	91.1	390	11	US-09-214-592-28	Sequence 28, Appl
15	624	91.1	390	11	US-09-214-592-29	Sequence 29, Appl
16	624	91.1	390	11	US-09-214-592-33	Sequence 33, Appl
17	624	91.1	390	12	US-10-276-947-1	Sequence 1, Appli
18	624	91.1	390	15	US-10-087-268-2	Sequence 2, Appli
19	624	91.1	390	15	US-10-087-268-5	Sequence 5, Appli
20	624	91.1	391	11	US-09-214-592-17	Sequence 17, Appl
21	621	90.7	390	11	US-09-214-592-20	Sequence 20, Appl
22	621	90.7	390	11	US-09-214-592-23	Sequence 23, Appl
23	562	82.0	98	12	US-10-187-394-1	Sequence 1, Appli
24	540	78.8	116	14	US-10-115-406-24	Sequence 24, Appl
25	540	78.8	116	15	US-10-154-333-26	Sequence 26, Appl
26	540	78.8	373	11	US-09-214-592-32	Sequence 32, Appl
27	529	77.2	304	10	US-09-756-283A-26	Sequence 26, Appl
28	521	76.1	112	10	US-09-813-271B-8	Sequence 8, Appli
29	499	72.8	98	12	US-10-187-394-4	Sequence 4, Appli
30	491	71.7	114	14	US-10-115-406-25	Sequence 25, Appl
31	491	71.7	114	15	US-10-154-333-27	Sequence 27, Appl
32	491	71.7	382	11	US-09-214-592-34	Sequence 34, Appl
33	490	71.5	112	10	US-09-813-271B-12	Sequence 12, Appl
34	489	71.4	383	10	US-09-756-283A-27	Sequence 27, Appl
35	486	70.9	112	10	US-09-813-271B-6	Sequence 6, Appli
36	486	70.9	113	10	US-09-813-398-15	Sequence 15, Appl
37	486	70.9	114	10	US-09-813-459-24	Sequence 24, Appl
38	486	70.9	114	14	US-10-115-406-23	Sequence 23, Appl
39	486	70.9	114	15	US-10-154-333-25	Sequence 25, Appl
40	486	70.9	115	10	US-09-880-708-27	Sequence 27, Appl
41	486	70.9	410	11	US-09-214-592-22	Sequence 22, Appl
42	486	70.9	412	11	US-09-214-592-19	Sequence 19, Appl
43	486	70.9	412	14	US-10-028-158-21	Sequence 21, Appl
44	485	70.8	412	11	US-09-214-592-24	Sequence 24, Appl
45	483	70.5	412	11	US-09-214-592-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1

US-09-813-271B-2

; Sequence 2, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

; APPLICANT:

; (A) Nico Cerletti

; TITLE OF INVENTION: New process for the production of

; biologically active protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. US20020115834A1artis Patent Department

; STREET: 564 Morris Avenue

; CITY: Summit

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07901

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/813,271B

; FILING DATE: 20-Mar-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP95/02719

; FILING DATE: 12-Jul-95

; APPLICATION NUMBER: EPO 94810439.3

; FILING DATE: 25-Jul-94

; ATTORNEY/AGENT INFORMATION:

; NAME: Pfeiffer, Hesna J. .

; REGISTRATION NUMBER: 22640

; REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (908) 522-6940

; TELEFAX: (908) 522-6955

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 112 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-813-271B-2

Query Match 91.1%; Score 624; DB 10; Length 112;

Best Local Similarity 93.3%; Pred. No. 1.5e-57;

Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60

||||||| |||||||||

Db 1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52

```

Qy      61 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      53 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

```

RESULT 2

```

US-09-813-398-13
; Sequence 13, Application US/09813398
; Patent No. US20020169292A1
; GENERAL INFORMATION:
; APPLICANT: Bruce D. Weintraub
; APPLICANT: Mariusz W. Szkudlinski
; APPLICANT: University of Maryland
; TITLE OF INVENTION: CYSTINE KNOT GROWTH FACTOR MUTANTS
; FILE REFERENCE: UOFMD.003C1
; CURRENT APPLICATION NUMBER: US/09/813,398
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: PCT/US99/05908
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/US98/19772
; PRIOR FILING DATE: 1998-09-22
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 113
; TYPE: PRT
; ORGANISM: HOMO SAPIEN
US-09-813-398-13

```

```

Query Match          91.1%; Score 624; DB 10; Length 113;
Best Local Similarity 93.3%; Pred. No. 1.5e-57;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
          ||||||||||| |||||||||||||||||||||||||||||||||||||||
Db      2 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 53

```

```

Qy      61 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 113

```

RESULT 3

```

US-09-813-459-22
; Sequence 22, Application US/09813459
; Patent No. US20020107369A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Se-Jin
; Cunningham, No. US20020107369Aleen
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-10
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA

```

```

;           ZIP: 90067
;
;   COMPUTER READABLE FORM:
;
;           MEDIUM TYPE: Floppy disk
;
;           COMPUTER: IBM PC compatible
;
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;
;           SOFTWARE: PatentIn Release #1.0, Version #1.25
;
;   CURRENT APPLICATION DATA:
;
;           APPLICATION NUMBER: US/09/813,459
;
;           FILING DATE: 20-Mar-2001
;
;           CLASSIFICATION: <Unknown>
;
;   PRIOR APPLICATION DATA:
;
;           APPLICATION NUMBER: 08/624,635
;
;           FILING DATE: <Unknown>
;
;   ATTORNEY/AGENT INFORMATION:
;
;           NAME: Wetherell, Jr., Ph.D., John R.,
;
;           REGISTRATION NUMBER: 31,678
;
;           REFERENCE/DOCKET NUMBER: PD-3054
;
;   TELECOMMUNICATION INFORMATION:
;
;           TELEPHONE: (619) 455-5100
;
;           TELEFAX: (619) 455-5110
;
;   INFORMATION FOR SEQ ID NO: 22:
;
;   SEQUENCE CHARACTERISTICS:
;
;           LENGTH: 114 amino acids
;
;           TYPE: amino acid
;
;           STRANDEDNESS: single
;
;           TOPOLOGY: linear
;
;   MOLECULE TYPE: protein
;
;   IMMEDIATE SOURCE:
;
;           CLONE: TGF-beta-1
;
;   FEATURE:
;
;           NAME/KEY: Protein
;
;           LOCATION: 1..114
;
;   SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-813-459-22

```

```

Query Match          91.1%;  Score 624;  DB 10;  Length 114;
Best Local Similarity 93.3%;  Pred. No. 1.5e-57;
Matches 112;  Conservative 0;  Mismatches 0;  Indels 8;  Gaps 1;

```

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
        |||||||||  |||||||||||||||||||||||||||||||||||
Db      3 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 54

Qy      61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 120
        |||||||||||||||||||||||||||||||||||||||||||||||
Db      55 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 114

```

```

RESULT 4
US-10-115-406-21
; Sequence 21, Application US/10115406
; Publication No. US20020127612A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: LEE, Se-Jin
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-9
; FILE REFERENCE: JHU1190-3

```

```

; CURRENT APPLICATION NUMBER: US/10/115,406
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/301,520
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: US 09/172,062
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 08/491,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: PCT/US94/00685
; PRIOR FILING DATE: 1994-01-12
; PRIOR APPLICATION NUMBER: US 08/003,303
; PRIOR FILING DATE: 1993-01-12
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-115-406-21

```

```

Query Match          91.1%; Score 624; DB 14; Length 114;
Best Local Similarity 93.3%; Pred. No. 1.5e-57;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
        |||
Db      3 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 54
        |||

Qy     61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
        |||
Db     55 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 114
        |||

```

RESULT 5

US-10-154-333-23

```

; Sequence 23, Application US/10154333
; Publication No. US20030109684A1
; GENERAL INFORMATION:
; APPLICANT: JOHNS HOPKINS UNIVERSITY
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-3
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: US
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/154,333
; FILING DATE: 21-May-2002
; CLASSIFICATION: <Unknown>

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/389,705
; FILING DATE: 03-Sep-1999
; APPLICATION NUMBER: 09/153,733
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: WETHERELL, JR. Ph.D., JOHN R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: FD2279 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: TGF-beta 1
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..114
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-154-333-23

```

```

Query Match          91.1%;  Score 624;  DB 15;  Length 114;
Best Local Similarity 93.3%;  Pred. No. 1.5e-57;
Matches 112;  Conservative 0;  Mismatches 0;  Indels 8;  Gaps 1;

```

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
        |||||||||  |||||||||||||||||||||||||||||||||||||||
Db      3 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 54

Qy     61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSC KCS 120
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db     55 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSC KCS 114

```

```

RESULT 6
US-09-859-211-47
; Sequence 47, Application US/09859211
; Patent No. US20020157125A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Se-Jin
; APPLICANT: McPherron, Alexandra C.
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-8
; FILE REFERENCE: 07265/144001
; CURRENT APPLICATION NUMBER: US/09/859,211
; CURRENT FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 09/019,070
; PRIOR FILING DATE: 1998-02-05
; PRIOR APPLICATION NUMBER: 08/862,445
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 08/847,910
; PRIOR FILING DATE: 1997-04-28

```



```

;          FILING DATE: 12-JAN-1993
;  ATTORNEY/AGENT INFORMATION:
;          NAME: Lisa A. Haile, Ph.D.
;          REGISTRATION NUMBER: 38,347
;          REFERENCE/DOCKET NUMBER: 07265/057002
;  TELECOMMUNICATION INFORMATION:
;          TELEPHONE: 858/677-1456
;          TELEFAX: 619/677-1465
;  INFORMATION FOR SEQ ID NO: 25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 115 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;    MOLECULE TYPE: protein
;    IMMEDIATE SOURCE:
;      CLONE: TGF-beta-1
;    SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-880-708-25

```

```

Query Match          91.1%;  Score 624;  DB 10;  Length 115;
Best Local Similarity 93.3%;  Pred. No. 1.5e-57;
Matches 112;  Conservative 0;  Mismatches 0;  Indels 8;  Gaps 1;

```

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
        |||||||||  |||||||||||||||||||||||||||||||||||
Db      4 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 55

Qy     61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
        ||||||||||||||||||||||||||||||||||||||||||||
Db     56 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 115

```

RESULT 8

```

US-09-872-856-47
; Sequence 47, Application US/09872856
; Publication No. US20030074680A1
; GENERAL INFORMATION:
; APPLICANT: Johns Hopkins University School of Medicine
; APPLICANT: Lee, Se-Jin
; APPLICANT: McPherron, Alexandra
; TITLE OF INVENTION: Growth Differentiation Factor-8
; FILE REFERENCE: JHU1120-17
; CURRENT APPLICATION NUMBER: US/09/872,856
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 09/124,180
; PRIOR FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: US 09/019,070
; PRIOR FILING DATE: 1998-02-05
; PRIOR APPLICATION NUMBER: US 08/862,445
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 08/847,910
; PRIOR FILING DATE: 1997-04-28
; PRIOR APPLICATION NUMBER: US 08/795,071
; PRIOR FILING DATE: 1997-02-05
; PRIOR APPLICATION NUMBER: US 08/525,596
; PRIOR FILING DATE: 1995-10-25
; PRIOR APPLICATION NUMBER: PCT/US 94/03019

```

```
; PRIOR FILING DATE: 1994-03-18
; PRIOR APPLICATION NUMBER: US 08/033,923
; PRIOR FILING DATE: 1993-03-19
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
;   LENGTH: 115
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-872-856-47
```

RESULT 9

```

;          NAME: Wetherell, Jr., Ph.D, John R.
;          REGISTRATION NUMBER: 31,678
;          REFERENCE/DOCKET NUMBER: 07265/075001
;          TELECOMMUNICATION INFORMATION:
;          TELEPHONE: 619-678-5070
;          TELEFAX: 619-678-5099
;          INFORMATION FOR SEQ ID NO: 29:
;          SEQUENCE CHARACTERISTICS:
;          LENGTH: 115 amino acids
;          TYPE: amino acid
;          TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          IMMEDIATE SOURCE:
;          CLONE: TGF-beta-1
;          FEATURE:
;          NAME/KEY: Protein
;          LOCATION: 1..115
;          SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-10-335-483-29

```

```

Query Match          91.1%;  Score 624;  DB 15;  Length 115;
Best Local Similarity 93.3%;  Pred. No. 1.5e-57;
Matches 112;  Conservative 0;  Mismatches 0;  Indels 8;  Gaps 1;

```

```

Qy          1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
  |||||  |||||
Db          4 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 55

Qy          61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 120
  |||||  |||||
Db          56 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 115

```

RESULT 10

US-10-002-278-8

```

; Sequence 8, Application US/10002278
; Publication No. US20020132334A1
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M.
; APPLICANT: Basler, Konard
; APPLICANT: Yamada, Toshiya
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF DORSALIN-1
; FILE REFERENCE: 0575/40314-A
; CURRENT APPLICATION NUMBER: US/10/002,278
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: COOH-terminus of TGF-Beta 1
; NAME/KEY: DOMAIN
; LOCATION: (1)..(139)
; OTHER INFORMATION:
US-10-002-278-8

```

Query Match 91.1%; Score 624; DB 14; Length 139;
Best Local Similarity 93.3%; Pred. No. 1.9e-57;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
          |||
Db      28 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 79
          |||

Qy      61 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
          |||
Db      80 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 139
          |||
```

RESULT 11

US-09-214-592-25

; Sequence 25, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

; APPLICANT: Yamasaki, CMotoo

; APPLICANT: Shibata, CKenji

; APPLICANT: Sato, CYasufumi

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND

METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

; FILE REFERENCE: 11060

; CURRENT APPLICATION NUMBER: US/09/214,592A

; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE:

; SEQ ID NO 25

; LENGTH: 315

; TYPE: PRT

; ORGANISM: bovine

US-09-214-592-25

Query Match 91.1%; Score 624; DB 11; Length 315;
Best Local Similarity 93.3%; Pred. No. 4.5e-57;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
          |||
Db      204 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 255
          |||

Qy      61 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
          |||
Db      256 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 315
          |||
```

RESULT 12

US-09-756-283A-23

; Sequence 23, Application US/09756283A

; Patent No. US20020151478A1

; GENERAL INFORMATION:

; APPLICANT: Chernajovsky, Yuti

; APPLICANT: Dreja, Hanna Stina

; APPLICANT: Adams, Gillian

; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-756-283A-23

Query Match 91.1%; Score 624; DB 10; Length 390;
Best Local Similarity 93.3%; Pred. No. 5.6e-57;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      |||
Db      279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

Qy      61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      |||
Db      331 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 13

US-09-214-592-26

; Sequence 26, Application US/09214592A
; Publication No. US20030027218A1
; GENERAL INFORMATION:
; APPLICANT: Yamasaki,CMotoo
; APPLICANT: Shibata,CKenji
; APPLICANT: Sato,CYasufumi
; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND
METHOD
; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS
; FILE REFERENCE: 11060
; CURRENT APPLICATION NUMBER: US/09/214,592A
; CURRENT FILING DATE: 1999-01-18
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE:
; SEQ ID NO 26
; LENGTH: 390
; TYPE: PRT
; ORGANISM: porcine
US-09-214-592-26

Query Match 91.1%; Score 624; DB 11; Length 390;
Best Local Similarity 93.3%; Pred. No. 5.6e-57;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      |||
Db      279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

Qy      61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      |||
```

Db 331 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 14

US-09-214-592-28

; Sequence 28, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

; APPLICANT: Yamasaki, CMotoo

; APPLICANT: Shibata,CKenji

; APPLICANT: Sato, CYasufumi

; TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

; FILE REFERENCE: 11060

; CURRENT APPLICATION NUMBER: US/09/214,592A

; CURRENT FILING DATE: 1999-01-18

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE:

; SEQ ID NO 28

; LENGTH: 390

```
; TYPE: PRT
```

; ORGANISM: canine

US-09-214-592-28

Query Match 91.1%; Score 624; DB 11; Length 390;

Best Local Similarity 93.3%; Pred. No. 5.6e-57;

Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

Db 279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

Qy 61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSC KCS 120

[illegible]

Db 331 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 15

US-09-214-592-29

; Sequence 29, Application US/09214592A

; Publication No. US20030027218A1

; GENERAL INFORMATION:

APPLICANT: Yamasaki, CMotoo

; APPLICANT: Shibata,CKenji

; APPLICANT: Sato, CYasufumi

10 TITLE OF INVENTION: PEPTIDES WHICH PROMOTE ACTIVATION OF LATENT TGF- AND

METHOD

; TITLE OF INVENTION: OF SCREENING TGF- ACTIVITY-REGULATING COMPOUNDS

FILE REFERENCE: 11060

CURRENT APPLICATION NUMBER: US/09/214,592A

; CURRENT FILING DATE: 1999-01-18

NUMBER OF SEQ ID NOS: 34

; SOFTWARE:

; SEQ ID NO 29

LENGTH: 390

```

; TYPE: PRT

```

US-09-214-592-29

Query Match 91.1%; Score 624; DB 11; Length 390;
Best Local Similarity 93.3%; Pred. No. 5.6e-57;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 ||||||| |
Db 279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

Qy 61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPI VYYYVGRKP KVEQLSNMIVRSCKCS 120
 ||||| |
Db 331 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPI VYYYVGRKP KVEQLSNMIVRSCKCS 390

Search completed: October 28, 2003, 09:28:01
Job time : 8.28061 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

```
Run on:      October 28, 2003, 00:37:16 ; Search time 3.96533 Seconds
              (without alignments)
              2910.285 Million cell updates/sec
```

Title: US-10-017-372E-15
Perfect score: 685
Sequence: 1 ALDTNYCFSSTDYKDDDDKE.....GRKPKVEQLSNMIVRSCKCS 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :      PIR_76:*
            1:  pir1:*
            2:  pir2:*
            3:  pir3:*
            4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	624	91.1	315	2	A40057	transforming growth	
2	624	91.1	390	1	WFHU2	transforming growth	
3	624	91.1	390	2	JC4023	transforming growth	
4	624	91.1	390	2	A26960	transforming growth	
5	624	91.1	390	2	A27512	transforming growth	
6	624	91.1	390	2	I46463	transforming growth	
7	624	91.1	391	2	S01413	transforming growth	
8	621	90.7	390	1	WFMS2	transforming growth	
9	621	90.7	390	2	S10219	transforming growth	
10	609	88.9	130	2	I48196	transforming growth	
11	540	78.8	373	2	A41918	transforming growth	
12	491	71.7	382	2	B61036	transforming growth	
13	486	70.9	410	2	A41397	transforming growth	
14	486	70.9	412	2	A36169	transforming growth	
15	485	70.8	410	2	A55706	transforming growth	
16	483	70.5	412	2	A34939	transforming growth	
17	479	69.9	409	2	S01825	transforming growth	
18	475	69.3	412	2	A39489	transforming growth	
19	468	68.3	112	2	A61439	transforming growth	
20	468	68.3	414	1	WFMKB2	transforming growth	
21	468	68.3	414	2	A31249	transforming growth	
22	468	68.3	442	2	B31249	transforming growth	
23	466	68.0	414	1	WFMSB2	transforming growth	
24	464	67.7	413	1	WFXLB2	transforming growth	
25	209	30.5	425	2	I47072	inhibin beta-A cha	
26	205	29.9	424	1	WFPGBA	inhibin beta-A cha	
27	205	29.9	424	1	S31440	inhibin beta-A cha	
28	205	29.9	425	1	S50898	inhibin beta-A cha	
29	205	29.9	426	1	B24248	inhibin beta-A cha	
30	201	29.3	424	1	B40905	inhibin beta-A cha	
31	186	27.2	413	2	JC4862	activin beta-A cha	
32	185	27.0	115	2	PN0504	activin beta A cha	
33	185	27.0	350	2	JC5241	activin beta E cha	
34	178.5	26.1	207	2	S37618	vgr protein - rat	
35	178.5	26.1	513	1	BMHU6	bone morphogenetic	
36	178	26.0	455	2	A43918	TGF-beta-related p	
37	177	25.8	367	2	JC4151	activin beta D cha	
38	176.5	25.8	510	2	A54798	Vg-1-related prote	
39	169	24.7	352	2	JC2466	inhibin beta-C cha	
40	169	24.7	372	2	C39364	GDF-1 embryonic gr	
41	169	24.7	430	2	JQ1184	osteogenic protein	
42	168.5	24.6	349	1	WFPGBB	inhibin beta-B cha	
43	168.5	24.6	407	1	A40150	inhibin beta-B cha	
44	168.5	24.6	408	2	S50899	betaB inhibin prec	
45	167.5	24.5	370	2	I51199	activin beta B sub	

ALIGNMENTS

RESULT 1
A40057

transforming growth factor beta-1 precursor - bovine (fragment)

N;Alternate names: beta-TGF; cartilage-inducing factor-A; EGF-dependent TGF or dEGF; MGF-b; milk growth factor b; TGF-type II

C;Species: Bos primigenius taurus (cattle)

C;Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 16-Jul-1999

C;Accession: A40057; A42320; A05284; A24322; B61439

R;Van Obberghen-Schilling, E.; Kondaiah, P.; Ludwig, R.L.; Sporn, M.B.; Baker, C.C.

Mol. Endocrinol. 1, 693-698, 1987

A;Title: Complementary deoxyribonucleic acid cloning of bovine transforming growth factor-beta1.

A;Reference number: A40057; MUID:91042552; PMID:3153459

A;Accession: A40057

A;Molecule type: mRNA

A;Residues: 1-315 <VAN>

A;Cross-references: GB:M36271; NID:g163747; PIDN:AAA30778.1; PID:g163748

R;Ogawa, Y.; Schmidt, D.K.; Dasch, J.R.; Chang, R.J.; Glaser, C.B.

J. Biol. Chem. 267, 2325-2328, 1992

A;Title: Purification and characterization of transforming growth factor-beta2.3 and -beta1.2 heterodimers from bovine bone.

A;Reference number: A42320; MUID:92129307; PMID:1733936

A;Accession: A42320

A;Molecule type: protein

A;Residues: 204-209,'X',211-217 <OGA>

R;Roberts, A.B.; Anzano, M.A.; Meyers, C.A.; Wideman, J.; Blacher, R.; Pan, Y.C.E.; Stein, S.; Lehrman, S.R.; Smith, J.M.; Lamb, L.C.; Sporn, M.B.

Biochemistry 22, 5692-5698, 1983

A;Title: Purification and properties of a type beta transforming growth factor from bovine kidney.

A;Reference number: A05284; MUID:84104793; PMID:6607069

A;Accession: A05284

A;Molecule type: protein

A;Residues: 204-218 <ROB>

R;Seyedin, S.M.; Thompson, A.Y.; Bentz, H.; Rosen, D.M.; McPherson, J.M.; Conti, A.; Siegel, N.R.; Galluppi, G.R.; Piez, K.A.

J. Biol. Chem. 261, 5693-5695, 1986

A;Title: Cartilage-inducing factor-A. Apparent identity to transforming growth factor-beta.

A;Reference number: A24322; MUID:86195954; PMID:3754555

A;Accession: A24322

A;Molecule type: protein

A;Residues: 204-233 <SEY>

R;Jin, Y.; Cox, D.A.; Knecht, R.; Raschdorf, F.; Cerletti, N.

J. Protein Chem. 10, 565-575, 1991

A;Title: Separation, purification, and sequence identification of TGF-beta1 and TGF-beta2 from bovine milk.

A;Reference number: A61439; MUID:92189724; PMID:1799413

A;Accession: B61439

A;Molecule type: protein

A;Residues: 204-209,'X',211-217,'XX',220-232 <JIN>

C;Comment: This polypeptide is composed of two polypeptide chains cross-linked by disulfide bonds. It has been found in neoplastic and non-neoplastic tissues.

C;Comment: Type II TGF does not bind to the EGF receptor and lacks intrinsic mitogenic activity, but in soft agar, it reacts synergistically with either type I TFG or EGF, and induces cell proliferation. Cells grown in monolayer do not respond in a similar manner to these growth factors, but morphologically do acquire a transformed phenotype.

C;Superfamily: inhibin
C;Keywords: glycoprotein; growth factor; heterodimer
F;204-315/Product: transforming growth factor beta-1 #status experimental <MAT>
F;7,61,101/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.1%; Score 624; DB 2; Length 315;
Best Local Similarity 93.3%; Pred. No. 5.9e-58;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
          |||||
Db      204 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 255

Qy      61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
          |||||
Db      256 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 315
```

RESULT 2

WFHU2

transforming growth factor beta-1 precursor [validated] - human

N;Alternate names: growth-inhibitory factor; TGF type 2; TGF-beta

C;Species: Homo sapiens (man)

C;Date: 28-Feb-1986 #sequence_revision 19-Oct-1995 #text_change 08-Dec-2000

C;Accession: A27513; A01395; A22290; I59664; S53444

R;Derynck, R.; Rhee, L.; Chen, E.Y.; Van Tilburg, A.

Nucleic Acids Res. 15, 3188-3189, 1987

A;Title: Intron-exon structure of the human transforming growth factor-beta precursor gene.

A;Reference number: A27513; MUID:87174845; PMID:3470709

A;Accession: A27513

A;Molecule type: DNA

A;Residues: 1-390 <DER>

A;Cross-references: GB:X05839; GB:Y00112; NID:g37097; PIDN:CAA29283.1;
PID:g1212989

R;Derynck, R.; Jarrett, J.A.; Chen, E.Y.; Eaton, D.H.; Bell, J.R.; Assoian, R.K.; Roberts, A.B.; Sporn, M.B.; Goeddel, D.V.

Nature 316, 701-705, 1985

A;Title: Human transforming growth factor-beta complementary DNA sequence and expression in normal and transformed cells.

A;Reference number: A01395; MUID:85296301; PMID:3861940

A;Accession: A01395

A;Molecule type: mRNA

A;Residues: 1-9,'P',11-24,'P',26-159,'R',160-390 <DE2>

A;Cross-references: GB:X02812; GB:J05114; NID:g37092; PIDN:CAA26580.1;
PID:g37093

A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide

R;Massague, J.; Like, B.

J. Biol. Chem. 260, 2636-2645, 1985

A;Title: Cellular receptors for type beta transforming growth factor. Ligand binding and affinity labeling in human and rodent cell lines.

A;Reference number: A22290; MUID:85131019; PMID:2982829

A;Accession: A22290

A;Molecule type: protein

A;Residues: 279-295,'XX',298-301 <MAS>

|||||
Db 331 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 5

A27512

transforming growth factor beta-1 precursor - pig

N;Alternate names: TGF-beta

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 16-Jul-1999

C;Accession: A27512; A26356; I46657

R;Derynck, R.; Rhee, L.

Nucleic Acids Res. 15, 3187, 1987

A;Title: Sequence of the porcine transforming growth factor-beta precursor.

A;Reference number: A27512; MUID:87174844; PMID:3470708

A;Accession: A27512

A;Molecule type: mRNA

A;Residues: 1-390 <DER>

R;Cheifetz, S.; Weatherbee, J.A.; Tsang, M.L.S.; Anderson, J.K.; Mole, J.E.;

Lucas, R.; Massague, J.

Cell 48, 409-415, 1987

A;Title: The transforming growth factor-beta system, a complex pattern of cross-reactive ligands and receptors.

A;Reference number: A90890; MUID:87102890; PMID:2879635

A;Accession: A26356

A;Molecule type: protein

A;Residues: 279-322 <CHE>

R;Kondaiah, P.; Van Obberghen-Schilling, E.; Ludwig, R.L.; Dhar, R.; Sporn, M.B.; Roberts, A.B.

J. Biol. Chem. 263, 18313-18317, 1988

A;Title: cDNA cloning of porcine transforming growth factor-beta 1 mRNAs.

Evidence for alternate splicing and polyadenylation.

A;Reference number: I46657; MUID:89054010; PMID:2461367

A;Accession: I46657

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-390 <KON>

A;Cross-references: GB:M23703; NID:g755044; PIDN:AAA64616.1; PID:g755045

C;Genetics:

A;Gene: TGFB; TGF-beta-1

C;Superfamily: inhibin

C;Keywords: growth factor

Query Match 91.1%; Score 624; DB 2; Length 390;

Best Local Similarity 93.3%; Pred. No. 7.4e-58;

Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60

|||||

Db 279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

Qy 61 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120

|||||

Db 331 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 6

I46463

transforming growth factor beta-1 - sheep

C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C;Date: 19-Dec-1997 #sequence_revision 19-Dec-1997 #text_change 24-Nov-1999

C;Accession: I46463; S45115

R;Woodall, C.J.; McLaren, L.J.; Watt, N.J.

Gene 150, 371-373, 1994

A;Title: Sequence and chromosomal localisation of the gene encoding ovine latent transforming growth factor-beta 1.

A;Reference number: I46463; MUID:95121932; PMID:7821809

A;Accession: I46463

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-390 <WOO>

A;Cross-references: EMBL:X76916; NID:g496648; PIDN:CAA54242.1; PID:g496649

A;Note: submitted to the EMBL Data Library, December 1993

C;Superfamily: inhibin

Query Match 91.1%; Score 624; DB 2; Length 390;
Best Local Similarity 93.3%; Pred. No. 7.4e-58;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
          |||
Db      279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

Qy      61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 120
          |||
Db      331 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 7

S01413

transforming growth factor beta-1 precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 24-Nov-1999

C;Accession: S01413

R;Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.

Nucleic Acids Res. 16, 8730, 1988

A;Title: Nucleotide sequence of chicken transforming growth factor-beta 1 (TGF-beta 1).

A;Reference number: S01413; MUID:88335639; PMID:3166520

A;Accession: S01413

A;Molecule type: DNA

A;Residues: 1-391 <JAK>

A;Cross-references: EMBL:X12373; NID:g63808; PIDN:CAA30933.1; PID:g63809

C;Superfamily: inhibin

C;Keywords: growth factor

Query Match 91.1%; Score 624; DB 2; Length 391;
Best Local Similarity 93.3%; Pred. No. 7.4e-58;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
          |||
Db      280 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 331
```

```

Qy      61 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      |||
Db      332 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 391

```

RESULT 8

WFMS2

transforming growth factor beta-1 precursor - mouse

N;Alternate names: TGF type 2; TGF-beta

C;Species: Mus musculus (house mouse)

C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 24-Nov-1999

C;Accession: A01396

R;Derynck, R.; Jarrett, J.A.; Chen, E.Y.; Goeddel, D.V.

J. Biol. Chem. 261, 4377-4379, 1986

A;Title: The murine transforming growth factor-beta precursor.

A;Reference number: A01396; MUID:86168129; PMID:3007454

A;Accession: A01396

A;Molecule type: mRNA

A;Residues: 1-390 <DER>

A;Cross-references: GB:M13177; NID:g201952; PIDN:AAA40423.1; PID:g201953

A;Note: the authors suggest that residues 8-23 could represent the hydrophobic core of an amino-terminal signal peptide

C;Comment: The mature protein is the carboxyl-terminal segment of a precursor polypeptide; the active molecule is a dimer of identical polypeptide chains linked by an interchain disulfide bond.

C;Superfamily: inhibin

C;Keywords: glycoprotein; growth factor; growth regulation; homodimer; mitogen; transformation

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-278/Domain: propeptide #status predicted <PRO>

F;244-246/Region: cell attachment (R-G-D) motif

F;279-390/Product: transforming growth factor beta-1 #status predicted <MAT>

F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match          90.7%;  Score 621;  DB 1;  Length 390;
Best Local Similarity 92.5%;  Pred. No. 1.5e-57;
Matches 111;  Conservative 1;  Mismatches 0;  Indels 8;  Gaps 1;

```

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      |||
Db      279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

```

```

Qy      61 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      |||
Db      331 SLDTQYSKVLALYNQHNPASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 9

S10219

transforming growth factor beta-1 precursor - rat

N;Alternate names: TGF type 2; TGF-beta

C;Species: Rattus norvegicus (Norway rat)

C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 24-Nov-1999

C;Accession: S10219; PT0023; S02267

R;Qian, S.W.; Kondaiah, P.; Roberts, A.B.; Sporn, M.B.

Nucleic Acids Res. 18, 3059, 1990

A;Title: cDNA cloning by PCR of rat transforming growth factor beta-1.

A;Reference number: S10219; MUID:90272425; PMID:2349108
A;Accession: S10219
A;Molecule type: mRNA
A;Residues: 1-390 <QIA>
A;Cross-references: EMBL:X52498; NID:g57341; PIDN:CAA36741.1; PID:g57342
R;Okada, F.; Yamaguchi, K.; Ichihara, A.; Nakamura, T.
J. Biochem. 106, 304-310, 1989
A;Title: Purification and structural analysis of a latent form of transforming growth factor-beta from rat platelets.
A;Reference number: PT0023; MUID:90036779; PMID:2478527
A;Accession: PT0023
A;Molecule type: protein
A;Residues: 30-32,'X',34-38,'Q',40-42,'X',44 <OKA>
R;Okada, F.; Yamaguchi, K.; Ichihara, A.; Nakamura, T.
FEBS Lett. 242, 240-244, 1989
A;Title: One of two subunits of masking protein in latent TGF-beta is a part of pro-TGF-beta.
A;Reference number: S02267; MUID:89121078; PMID:2914605
A;Accession: S02267
A;Molecule type: protein
A;Residues: 30-32,'X',34-38,'Q',40-42,'X',44 <OK2>
C;Superfamily: inhibin
C;Keywords: glycoprotein; growth factor; integrin binding
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-278/Domain: propeptide #status experimental <PRO>
F;244-246/Region: cell attachment (R-G-D) motif
F;279-390/Product: transforming growth factor beta-1 #status predicted <MAT>
F;82,136,176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.7%; Score 621; DB 2; Length 390;
Best Local Similarity 92.5%; Pred. No. 1.5e-57;
Matches 111; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      |||
Db      279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

Qy      61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      |||
Db      331 SLDTQYSKVLALYNQHNP GASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 10

I48196
transforming growth factor beta-1 precursor - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 02-Jul-1996 #sequence_revision 04-Oct-1996 #text_change 24-Nov-1999
C;Accession: I48196
R;Wong, D.T.; Donoff, R.B.; Yang, J.; Song, B.Z.; Matossian, K.; Nagura, N.; Elovic, A.; McBride, J.; Gallagher, G.; Todd, R.; Chiang, T.; Chou, L.S.S.; Yung, C.M.; Galli, S.J.; Weller, P.F.
Am. J. Pathol. 143, 130-142, 1993
A;Title: Sequential expression of transforming growth factors alpha and beta 1 by eosinophils during cutaneous wound healing in the hamster.
A;Reference number: I48196; MUID:93304479; PMID:8317544
A;Accession: I48196
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA
A;Residues: 1-130 <RES>
A;Cross-references: EMBL:X60296; NID:g396177; PIDN:CAA42838.1; PID:g396178
C;Superfamily: inhibin

Query Match 88.9%; Score 609; DB 2; Length 130;
Best Local Similarity 91.7%; Pred. No. 8.7e-57;
Matches 110; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
          |||||
Db      19 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 70

Qy     61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
          |||||
Db     71 SLDTQYSKVLALYNQHNP GASAGPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSYKCS 130
```

RESULT 11

A41918

transforming growth factor beta-4 precursor - chicken (fragment)

N;Alternate names: TGF-beta 4

C;Species: Gallus gallus (chicken)

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999

C;Accession: A41918; A34941; S03110

R;Burt, D.W.; Jakowlew, S.B.

Mol. Endocrinol. 6, 989-992, 1992

A;Title: Correction: a new interpretation of a chicken transforming growth factor-beta 4 complementary DNA.

A;Reference number: A41918; MUID:92357039; PMID:1353860

A;Accession: A41918

A;Molecule type: mRNA

A;Residues: 1-373 <BUR>

A;Cross-references: GB:M31160; GB:X08012; GB:S41706; NID:g1262437;

PIDN:AAB05637.1; PID:g1262438

A;Note: sequence extracted from NCBI backbone (NCBIN:110186, NCBIP:110187)

A;Note: this report corrects and reinterprets the sequence from reference A34941

R;Jakowlew, S.B.; Dillard, P.J.; Sporn, M.B.; Roberts, A.B.

Mol. Endocrinol. 2, 1186-1195, 1988

A;Title: Complementary deoxyribonucleic acid cloning of a messenger ribonucleic acid encoding transforming growth factor beta-4 from chicken embryo chondrocytes.

A;Reference number: A34941; MUID:89112198; PMID:2464131

A;Accession: A34941

A;Molecule type: mRNA

A;Residues: 'MDPMSIGPKSCGGSPWRPPGTAPWSIGSRRATASSSCSTSSRVRAEVGGRAL',122-209,'D',211-373 <JAK>

A;Cross-references: EMBL:X08012

A;Note: this sequence has been corrected in A41918

C;Superfamily: inhibin

C;Keywords: glycoprotein; growth factor

F;1/Domain: signal sequence (fragment) #status predicted <SIG>

F;223-225/Region: cell attachment (R-G-D) motif

F;260-373/Product: transforming growth factor beta-4 #status predicted <MAT>

F;54,109,153/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 78.8%; Score 540; DB 2; Length 373;

Best Local Similarity 79.2%; Pred. No. 4.7e-49;
Matches 95; Conservative 8; Mismatches 9; Indels 8; Gaps 2;

```
Qy      2 LDTNYCFS-STDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      |||:|||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      261 LDTDYCFGPGTD-----EKNCCVRPLYIDFRKDLQWKWIHEPKGYMANFCMGPCPYIW 313

Qy      61 SLDTQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      | ||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      314 SADTQYTKVLALYNQHNPASAAPCCVPQTLDPPIIYYVGRNVRVEQLSNMVVRACKCS 373
```

RESULT 12

B61036

transforming growth factor beta-5 precursor - African clawed frog

C;Species: *Xenopus laevis* (African clawed frog)

C;Date: 31-Dec-1993 #sequence_revision 03-Feb-1994 #text_change 16-Jul-1999

C;Accession: A34929; B61036

R;Kondaiah, P.; Sands, M.J.; Smith, J.M.; Fields, A.; Roberts, A.B.; Sporn, M.B.; Melton, D.A.

J. Biol. Chem. 265, 1089-1093, 1990

A;Title: Identification of a novel transforming growth factor-beta (TGF-beta5) mRNA in *Xenopus laevis*.

A;Reference number: A34929; MUID:90110090; PMID:2295601

A;Accession: A34929

A;Molecule type: mRNA

A;Residues: 1-382 <KON>

A;Cross-references: GB:J05180; NID:g214821; PIDN:AAA49968.1; PID:g214822

R;Roberts, A.B.; Rosa, F.; Roche, N.S.; Coligan, J.E.; Garfield, M.; Rebbert, M.L.; Kondaiah, P.; Danielpour, D.; Kehrl, J.H.; Wahl, S.M.; Dawid, I.B.; Sporn, M.B.

Growth Factors 2, 135-147, 1990

A;Title: Isolation and characterization of TGF-beta2 and TGF-beta5 from medium conditioned by *Xenopus* XTC cells.

A;Reference number: A61036; MUID:90253806; PMID:2340184

A;Accession: B61036

A;Molecule type: protein

A;Residues: 271-276, 'X', 278-284, 'XX', 287-299 <ROB>

C;Superfamily: inhibin

C;Keywords: growth factor

F;271-382/Product: transforming growth factor beta-5 #status experimental <MAT>

Query Match 71.7%; Score 491; DB 2; Length 382;

Best Local Similarity 73.0%; Pred. No. 6.8e-44;

Matches 84; Conservative 10; Mismatches 13; Indels 8; Gaps 1;

```
Qy      6 YCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQ 65
      ||| : ||||: |||:||||| ||||| ||||| ||||| ||||| |||||
Db      276 YCFGNG-----PNCCVKPLYINFRKDLGWKWIHEPKGYEANYCLGNCPYIWSMDTQ 327

Qy      66 YSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      ||||:||||:||||: ||||| |||||:||||| ||||| ||||| |||||
Db      328 YSKVLSLYNQNNPGASISPCCVPDVLEPLPIIYYVGRTAKVEQLSNMVVRSCNCS 382
```

RESULT 13

A41397

transforming growth factor beta-3 precursor - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 16-Jul-1999
 C;Accession: A41397; A61039; A61225
 R;Miller, D.A.; Lee, A.; Matsui, Y.; Chen, E.Y.; Moses, H.L.; Derynck, R.
 Mol. Endocrinol. 3, 1926-1934, 1989
 A;Title: Complementary DNA cloning of the murine transforming growth factor-
 beta3 (TGFbeta3) precursor and the comparative expression of TGFbeta3 and
 TGFbeta1 messenger RNA in murine embryos and adult tissues.
 A;Reference number: A41397; MUID:90190650; PMID:2628730
 A;Accession: A41397
 A;Molecule type: mRNA
 A;Residues: 1-410 <MIL>
 A;Cross-references: GB:M32745; NID:g201949; PIDN:AAA40422.1; PID:g201950
 R;Denhez, F.; Lafyatis, R.; Kondaiah, P.; Roberts, A.B.; Sporn, M.B.
 Growth Factors 3, 139-146, 1990
 A;Title: Cloning by polymerase chain reaction of a new mouse TGF-beta, mTGF-
 beta3.
 A;Reference number: A61039; MUID:91000714; PMID:2206556
 A;Accession: A61039
 A;Molecule type: mRNA
 A;Residues: 1-410 <DEN>
 R;Watrin, F.; Scotto, L.; Assoian, R.K.; Wolgemuth, D.J.
 Cell Growth Differ. 2, 77-83, 1991
 A;Title: Cell lineage specificity of expression of the murine transforming
 growth factor beta-3 and transforming growth factor beta-1 genes.
 A;Reference number: A61225; MUID:91299576; PMID:2069871
 A;Accession: A61225
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 285-410 <WAT>
 C;Superfamily: inhibin
 C;Keywords: glycoprotein; growth factor; growth regulation
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-298/Domain: propeptide #status predicted <PRO>
 F;259-261/Region: cell attachment (R-G-D) motif
 F;299-410/Product: transforming growth factor beta-3 #status predicted <MAT>
 F;72,133,140/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.9%; Score 486; DB 2; Length 410;
 Best Local Similarity 71.7%; Pred. No. 2.5e-43;
 Matches 86; Conservative 11; Mismatches 15; Indels 8; Gaps 1;

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      ||||| : ||||| |||||:|||||:|||||:|||||:|||||
Db      299 ALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKWWHEPKGYANFCSGPCPYLR 350

Qy      61 SLDTQYSKVLALYNQHNPASAPCCVPOALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      ||| : || ||| || |||:||||| ||||| :||||| |||||:|:|||||
Db      351 SADTTHSTVLGLYNTLNPEASAPCCVPQDLEPLTILYVGRTPKVEQLSNMVVKCKCS 410

```

RESULT 14
 A36169
 transforming growth factor beta-3 precursor - human
 C;Species: Homo sapiens (man)
 C;Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 20-Jun-2000

C;Accession: A36169; A41262; S01824
R;ten Dijke, P.; Hansen, P.; Iwata, K.K.; Pieler, C.; Foulkes, J.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 4715-4719, 1988
A;Title: Identification of another member of the transforming growth factor type
beta gene family.
A;Reference number: A36169; MUID:88263019; PMID:3164476
A;Accession: A36169
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-412 <TEN>
A;Cross-references: GB:J03241; NID:g339551; PIDN:AAA61161.1; PID:g339552
R;Arrick, B.A.; Lee, A.L.; Grendell, R.L.; Derynck, R.
Mol. Cell. Biol. 11, 4306-4313, 1991
A;Title: Inhibition of translation of transforming growth factor-beta3 mRNA by
its 5' untranslated region.
A;Reference number: A41262; MUID:91342629; PMID:1875922
A;Accession: A41262
A;Molecule type: DNA
A;Residues: 1-48 <ARR>
A;Cross-references: GB:M58524
R;Derynck, R.; Lindquist, P.B.; Lee, A.; Wen, D.; Tamm, J.; Graycar, J.L.; Rhee,
L.; Mason, A.J.; Miller, D.A.; Coffey, R.J.; Moses, H.L.; Chen, E.Y.
EMBO J. 7, 3737-3743, 1988
A;Title: A new type of transforming growth factor-beta, TGF-beta3.
A;Reference number: S01824; MUID:89091120; PMID:3208746
A;Accession: S01824
A;Molecule type: mRNA
A;Residues: 3-412 <DER>
A;Cross-references: EMBL:X14885; NID:g37075; PIDN:CAA33024.1; PID:g1200236
C;Genetics:
A;Gene: GDB:TGFB3
A;Cross-references: GDB:120437; OMIM:190230
A;Map position: 14q24-14q24
A;Introns: 118/1; 172/3; 216/1; 252/1; 309/2; 360/3
C;Superfamily: inhibin
C;Keywords: growth factor; homodimer
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-300/Domain: propeptide #status predicted <PRO>
F;301-412/Product: transforming growth factor beta-3 #status predicted <MAT>

Query Match 70.9%; Score 486; DB 2; Length 412;
Best Local Similarity 71.7%; Pred. No. 2.5e-43;
Matches 86; Conservative 11; Mismatches 15; Indels 8; Gaps 1;

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      ||||| : ||||| |||||:|||||:|||||:|||||:|||||:
Db      301 ALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKWVHEPKGYANFCSGPCPYLR 352

Qy      61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      || :| || ||| || |||:||||| ||||| :||||| |||||:|:|||||
Db      353 SADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS 412

```

RESULT 15
A55706
transforming growth factor beta-3 precursor - rat
C;Species: Rattus norvegicus (Norway rat)

C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 16-Jul-1999
 C;Accession: A55706; B40699; S36042
 R;Wang, J.; Kuliszewski, M.; Yee, W.; Sedlackova, L.; Xu, J.; Tseu, I.; Post, M.
 J. Biol. Chem. 270, 2722-2728, 1995
 A;Title: Cloning and expression of glucocorticoid-induced genes in fetal rat
 lung fibroblasts. Transforming growth factor-beta-3.
 A;Reference number: A55706; MUID:95155340; PMID:7852342
 A;Accession: A55706
 A;Molecule type: mRNA
 A;Residues: 1-410 <WAN>
 A;Cross-references: GB:U03491
 A;Note: it is uncertain whether Met-1 is the initiator
 R;McKinnon, R.D.; Piras, G.; Ida Jr., J.A.; Dubois-Dalcq, M.
 J. Cell Biol. 121, 1397-1407, 1993
 A;Title: A role for TGF-beta in oligodendrocyte differentiation.
 A;Reference number: A40699; MUID:93286190; PMID:8509457
 A;Accession: B40699
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 157-211 <MCK>
 A;Cross-references: EMBL:X71903; NID:g311326; PIDN:CAA50722.1; PID:g311327
 C;Superfamily: inhibin

Query Match 70.8%; Score 485; DB 2; Length 410;
 Best Local Similarity 71.7%; Pred. No. 3.1e-43;
 Matches 86; Conservative 11; Mismatches 15; Indels 8; Gaps 1;

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      ||||| : ||||| |||||:|||||:|||||:|||||:
Db      299 ALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKVVHEPKGYANFCSGPCPYLR 350

Qy      61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
      || :| || ||| || |||:||||| |||| :||||| |||||:|:|||||
Db      351 SSDTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS 410
  
```

Search completed: October 28, 2003, 09:09:48
 Job time : 3.96533 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 00:08:41 ; Search time 2.4052 Seconds
 (without alignments)
 2346.251 Million cell updates/sec

Title: US-10-017-372E-15
 Perfect score: 685
 Sequence: 1 ALDTNYCFSSTDYKDDDDKE.....GRKPKVEQLSNMIVRSCKCS 120

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	624	91.1	315	1	TGF1_BOVIN	P18341	bos taurus
2	624	91.1	390	1	TGF1_CANFA	P54831	canis famil
3	624	91.1	390	1	TGF1_CERAE	P09533	cercopithec
4	624	91.1	390	1	TGF1_HUMAN	P01137	homo sapien
5	624	91.1	390	1	TGF1_PIG	P07200	sus scrofa
6	624	91.1	390	1	TGF1_SHEEP	P50414	ovis aries
7	621	90.7	390	1	TGF1_MOUSE	P04202	mus musculu
8	621	90.7	390	1	TGF1_RAT	P17246	rattus norv
9	620	90.5	390	1	TGF1_HORSE	O19011	equus cabal
10	612	89.3	390	1	TGF1_CAVPO	Q9z1y6	cavia porce
11	540	78.8	373	1	TGF1_CHICK	P09531	gallus gall
12	491	71.7	382	1	TGF1_XENLA	P16176	xenopus lae
13	486	70.9	410	1	TGF3_MOUSE	P17125	mus musculu
14	486	70.9	412	1	TGF3_HUMAN	P10600	homo sapien
15	485	70.8	412	1	TGF3_RAT	Q07258	rattus norv
16	479	69.9	409	1	TGF3_PIG	P15203	sus scrofa
17	475	69.3	412	1	TGF2_CHICK	P30371	gallus gall
18	468	68.3	112	1	TGF2_BOVIN	P21214	bos taurus
19	468	68.3	414	1	TGF2_HUMAN	P08112	homo sapien
20	468	68.3	435	1	TGF2_PIG	P09858	sus scrofa
21	466	68.0	414	1	TGF2_MOUSE	P27090	mus musculu
22	466	68.0	442	1	TGF2_RAT	Q07257	rattus norv
23	464	67.7	412	1	TGF3_CHICK	P16047	gallus gall
24	464	67.7	413	1	TGF2_XENLA	P17247	xenopus lae
25	209	30.5	425	1	IHBA_SHEEP	P43032	ovis aries
26	209	30.5	426	1	IHBA_HORSE	P55102	equus cabal
27	205	29.9	424	1	IHBA_MOUSE	Q04998	mus musculu
28	205	29.9	424	1	IHBA_PIG	P03970	sus scrofa
29	205	29.9	424	1	IHBA_RAT	P18331	rattus norv
30	205	29.9	425	1	IHBA_BOVIN	P07995	bos taurus
31	205	29.9	426	1	IHBA_HUMAN	P08476	homo sapien
32	202.5	29.6	374	1	GDF8_BRARE	O42222	brachydanio
33	196.5	28.7	375	1	GDF8_PAPHA	O18828	papio hamad
34	196	28.6	424	1	IHBA_CHICK	P27092	gallus gall
35	193.5	28.2	375	1	GDF8_SHEEP	O18830	ovis aries
36	192.5	28.1	375	1	GDF8_CHICK	O42220	gallus gall

37	192.5	28.1	375	1	GDF8_HUMAN	014793	homo sapien
38	192.5	28.1	375	1	GDF8_MELGA	042221	meleagris g
39	192.5	28.1	375	1	GDF8_PIG	018831	sus scrofa
40	192.5	28.1	376	1	GDF8_MOUSE	008689	mus musculu
41	192.5	28.1	376	1	GDF8_RAT	035312	rattus norv
42	190.5	27.8	405	1	GDFB_MOUSE	Q9zlw4	mus musculu
43	190.5	27.8	407	1	GDFB_HUMAN	095390	homo sapien
44	187.5	27.4	375	1	GDF8_BOVIN	018836	bos taurus
45	186	27.2	355	1	DVR1_BRARE	P35621	brachydanio

ALIGNMENTS

RESULT 1

TGF1_BOVIN

ID TGF1_BOVIN STANDARD; PRT; 315 AA.
AC P18341;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1) (Fragment).
GN TGFB1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91042552; PubMed=3153459;
RA van Obberghen-Schilling E., Kondaiah P., Ludwig R.L., Sporn M.B.,
RA Baker C.C.;
RT "Complementary deoxyribonucleic acid cloning of bovine transforming
RT growth factor-beta 1.";
RL Mol. Endocrinol. 1:693-698(1987).
RN [2]
RP SUBUNITS.
RC TISSUE=Bone;
RX MEDLINE=92129307; PubMed=1733936;
RA Ogawa Y., Schmidt D.K., Dasch J.R., Chang R.J., Glaser C.B.;
RT "Purification and characterization of transforming growth factor-beta
RT 2.3 and -beta 1.2 heterodimers from bovine bone.";
RL J. Biol. Chem. 267:2325-2328(1992).
CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC -!- SUBUNIT: Homodimer; disulfide-linked. Heterodimers of TGF-beta 1/2
CC have been found in bone.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the TGF-beta family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M36271; AAA30778.1; -.
 DR PIR; A40057; A40057.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGfb.
 DR InterPro; IPR001839; TGfb.
 DR InterPro; IPR001111; TGfb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGfb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGfb; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein.
 FT NON_TER 1 1
 FT PROPEP <1 203
 FT CHAIN 204 315 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 210 219 BY SIMILARITY.
 FT DISULFID 218 281 BY SIMILARITY.
 FT DISULFID 247 312 BY SIMILARITY.
 FT DISULFID 251 314 BY SIMILARITY.
 FT DISULFID 280 280 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 169 171 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 315 AA; 36269 MW; C2717A23D994E00E CRC64;

Query Match 91.1%; Score 624; DB 1; Length 315;
 Best Local Similarity 93.3%; Pred. No. 1.9e-59;
 Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 204 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 255
 Qy 61 SLDTQYSKVLALYNQHNPASAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 120
 |||||
 Db 256 SLDTQYSKVLALYNQHNPASAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 315

RESULT 2

TGF1_CANFA

ID TGF1_CANFA STANDARD; PRT; 390 AA.
 AC P54831;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).
 GN TGFB1.

OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Jugular vein endothelial;
 RX MEDLINE=95237630; PubMed=7721110;
 RA Manning A.M., Auchampach J.A., Drong R.F., Slightom J.L.;
 RT "Cloning of a canine cDNA homologous to the human transforming growth
 RT factor-beta 1-encoding gene.";
 RL Gene 155:307-308(1995).
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L34956; AAA51458.1; -.
 DR PIR; JC4023; JC4023.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 278 BY SIMILARITY.
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).

DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFbBETA.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT PROPEP 17 278
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 390 AA; 44356 MW; DFF63E2BAB44320E CRC64;

Query Match 91.1%; Score 624; DB 1; Length 390;
 Best Local Similarity 93.3%; Pred. No. 2.4e-59;
 Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330
 QY 61 SLDTQYSKVLALYNQHNPAGASAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
 |||||
 Db 331 SLDTQYSKVLALYNQHNPAGASAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 4

TGF1_HUMAN

ID TGF1_HUMAN STANDARD; PRT; 390 AA.
 AC P01137; Q9UCG4;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).
 GN TGFb1 OR TGFb.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87174845; PubMed=3470709;
 RA Derynck R., Rhee L., Chen E.Y., van Tilburg A.;
 RT "Intron-exon structure of the human transforming growth factor-beta

RT precursor gene.";
 RL Nucleic Acids Res. 15:3188-3189(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85296301; PubMed=3861940;
 RA Derynck R., Jarrett J.A., Chen E.Y., Eaton D.H., Bell J.R.,
 RA Assoian R.K., Roberts A.B., Sporn M.B., Goeddel D.V.;
 RT "Human transforming growth factor-beta complementary DNA sequence and
 RT expression in normal and transformed cells.";
 RL Nature 316:701-705(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Duodenum, and Eye;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 279-390 FROM N.A.
 RC TISSUE=Carcinoma;
 RA Urushizaki Y., Niitsu Y., Terui T., Koshida Y., Mahara K., Kohgo Y.,
 RA Urushizaki I., Takahashi Y., Ito H.;
 RT "Cloning and expression of the gene for human transforming growth
 RT factor-beta in Escherichia coli.";
 RL Tumor Res. 22:41-55(1987).
 RN [5]
 RP SEQUENCE OF 279-329.
 RC TISSUE=Bladder carcinoma;
 RX MEDLINE=93229900; PubMed=8471846;
 RA Bourdrel L., Lin C.-H., Lauren S.L., Elmore R.H., Sugarman B.J.,
 RA Hu S., Westcott K.R.;
 RT "Recombinant human transforming growth factor-beta 1: expression by
 RT Chinese hamster ovary cells, isolation, and characterization.";
 RL Protein Expr. Purif. 4:130-140(1993).
 RN [6]
 RP SEQUENCE OF 279-301.
 RX MEDLINE=85131019; PubMed=2982829;
 RA Massague J., Like B.;
 RT "Cellular receptors for type beta transforming growth factor. Ligand
 RT binding and affinity labeling in human and rodent cell lines.";

RL J. Biol. Chem. 260:2636-2645(1985).
 RN [7]
 RP STRUCTURE BY NMR OF 279-390.
 RX MEDLINE=93144319; PubMed=8424942;
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,
 RA Torchia D.A.;
 RT "Transforming growth factor beta 1: NMR signal assignments of the
 RT recombinant protein expressed and isotopically enriched using Chinese
 RT hamster ovary cells.";
 RL Biochemistry 32:1152-1163(1993).
 RN [8]
 RP STRUCTURE BY NMR OF 279-390.
 RX MEDLINE=93144320; PubMed=8424943;
 RA Archer S.J., Bax A., Roberts A.B., Sporn M.B., Ogawa Y., Piez K.A.,
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,
 RA Torchia D.A.;
 RT "Transforming growth factor beta 1: secondary structure as determined
 RT by heteronuclear magnetic resonance spectroscopy.";
 RL Biochemistry 32:1164-1171(1993).
 RN [9]
 RP STRUCTURE BY NMR OF 279-390.
 RX MEDLINE=96266150; PubMed=8679613;
 RA Hinck A.P., Archer S.J., Qian S.W., Roberts A.B., Sporn M.B.,
 RA Weatherbee J.A., Tsang M.L.-S., Lucas R., Zheng B.-L., Wenker J.,
 RA Torchia D.A.;
 RT "Transforming growth factor beta 1: three-dimensional structure in
 RT solution and comparison with the X-ray structure of transforming
 RT growth factor beta 2.";
 RL Biochemistry 35:8517-8534(1996).
 CC -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,
 CC DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY
 CC CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE
 CC SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A
 CC POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X05839; CAA29283.1; -.
 DR EMBL; X05840; CAA29283.1; JOINED.
 DR EMBL; X05843; CAA29283.1; JOINED.
 DR EMBL; X05844; CAA29283.1; JOINED.
 DR EMBL; X05849; CAA29283.1; JOINED.
 DR EMBL; X05850; CAA29283.1; JOINED.
 DR EMBL; X02812; CAA26580.1; ALT_SEQ.
 DR EMBL; BC001180; AAH01180.1; -.
 DR EMBL; BC000125; AAH00125.1; -.

DR EMBL; BC022242; AAH22242.1; -.
 DR EMBL; M38449; AAA36735.1; -.
 DR PIR; A27513; WFHU2.
 DR PDB; 1KLA; 17-AUG-96.
 DR PDB; 1KLC; 17-AUG-96.
 DR PDB; 1KLD; 17-AUG-96.
 DR Genew; HGNC:11766; TGFB1.
 DR MIM; 190180; -.
 DR GO; GO:0006916; P:anti-apoptosis; TAS.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal; 3D-structure.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 278
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294
 FT DISULFID 293 356
 FT DISULFID 322 387
 FT DISULFID 326 389
 FT DISULFID 355 355 INTERCHAIN.
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
 FT CONFLICT 10 10 L -> P (IN REF. 2).
 FT CONFLICT 159 159 R -> RR (IN REF. 2).
 FT STRAND 281 281
 FT TURN 282 287
 FT STRAND 294 296
 FT STRAND 300 300
 FT TURN 302 305
 FT STRAND 313 313
 FT STRAND 317 317
 FT STRAND 321 323
 FT HELIX 335 346
 FT TURN 348 349
 FT STRAND 358 370
 FT TURN 371 372
 FT STRAND 373 387
 SQ SEQUENCE 390 AA; 44341 MW; 75391614250288FE CRC64;

Query Match 91.1%; Score 624; DB 1; Length 390;
 Best Local Similarity 93.3%; Pred. No. 2.4e-59;
 Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 ||||||||| |||||||||||||||||||||
 Db 279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330


```

RT  "The transforming growth factor-beta system, a complex pattern of
RT  cross-reactive ligands and receptors.";
RL  Cell 48:409-415(1987).
CC  -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC        PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC        TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC        HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC        ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC        A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC  -!- SUBUNIT: Homodimer; disulfide-linked.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: Belongs to the TGF-beta family.
CC  -!- CAUTION: REF.3 SEQUENCE WHICH WAS SAID TO ORIGINATE FROM CHICKEN
CC        WHITE LEGHORN, SEEMS (REF.4) TO ORIGINATE FROM PIG.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; Y00111; CAA68291.1; -.
DR  EMBL; M23703; AAA64616.1; -.
DR  EMBL; X12373; CAA30933.1; -.
DR  EMBL; AF461808; AAL57902.1; -.
DR  PIR; A27512; A27512.
DR  HSSP; P01137; 1KLA.
DR  InterPro; IPR002400; GF_cysknot.
DR  InterPro; IPR003911; TGF_TGFB.
DR  InterPro; IPR001839; TGFB.
DR  InterPro; IPR001111; TGFB_N.
DR  Pfam; PF00019; TGF-beta; 1.
DR  Pfam; PF00688; TGFB_propeptide; 1.
DR  PRINTS; PR00438; GFCYSKNOT.
DR  PRINTS; PR01423; TGFBETA.
DR  ProDom; PD000357; TGFB; 1.
DR  SMART; SM00204; TGFB; 1.
DR  PROSITE; PS00250; TGF_BETA_1; 1.
KW  Growth factor; Mitogen; Glycoprotein; Signal; Polymorphism.
FT  SIGNAL          1      23      POTENTIAL.
FT  PROPEP          24     278
FT  CHAIN            279     390      TRANSFORMING GROWTH FACTOR BETA 1.
FT  DISULFID        285     294      BY SIMILARITY.
FT  DISULFID        293     356      BY SIMILARITY.
FT  DISULFID        322     387      BY SIMILARITY.
FT  DISULFID        326     389      BY SIMILARITY.
FT  DISULFID        355     355      INTERCHAIN (BY SIMILARITY).
FT  CARBOHYD         82     82      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD        136     136      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD        176     176      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  SITE            244     246      CELL ATTACHMENT SITE (POTENTIAL).
FT  VARIANT         114     114      L -> V.
FT  CONFLICT         6       7      LR -> PG (IN REF. 3).
FT  CONFLICT        180     180      R -> G (IN REF. 3).
FT  CONFLICT        237     237      N -> NA (IN REF. 3).

```

SQ SEQUENCE 390 AA; 44294 MW; A6E2C3659FC384E6 CRC64;

Query Match 91.1%; Score 624; DB 1; Length 390;
Best Local Similarity 93.3%; Pred. No. 2.4e-59;
Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

```
Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
          |||||
Db      279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

Qy      61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
          |||||
Db      331 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390
```

RESULT 6

TGF1_SHEEP

ID TGF1_SHEEP STANDARD; PRT: 390 AA.
AC P50414;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1).
GN TGFB1.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95121932; PubMed=7821809;
RA Woodall C.J., McLaren L.J., Watt N.J.;
RT "Sequence and chromosomal localisation of the gene encoding ovine
RT latent transforming growth factor-beta 1.";
RL Gene 150:371-373(1994).
RN [2]
RP SEQUENCE OF 281-390 FROM N.A.
RC STRAIN=Merino; TISSUE=Skin;
RX MEDLINE=95268698; PubMed=7749621;
RA Sutton R., Ward W.G., Raphael K.A., Cam G.R.;
RT "Growth factor expression in skin during wool follicle development.";
RL Comp. Biochem. Physiol. 110B:697-705(1995).
CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the TGF-beta family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86168129; PubMed=3007454;
 RA Derynck R., Jarrett J.A., Chen E.Y., Goeddel D.V.;
 RT "The murine transforming growth factor-beta precursor.";
 RL J. Biol. Chem. 261:4377-4379(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=96096545; PubMed=8522200;
 RA Guron C., Sudarshan C., Raghov R.;
 RT "Molecular organization of the gene encoding murine transforming
 RT growth factor beta 1.";
 RL Gene 165:325-326(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6, and NOD/LT; TISSUE=Spleen;
 RA Poirot L., Benoist C., Mathis D.;
 RT "Transforming growth factor-beta 1 sequence and expression: no
 RT difference between NOD/Lt and C57Bl/6 mouse strains.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
 CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
 CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M13177; AAA40423.1; -.
 DR EMBL; L42462; AAB00138.1; -.
 DR EMBL; L42456; AAB00138.1; JOINED.
 DR EMBL; L42457; AAB00138.1; JOINED.
 DR EMBL; L42458; AAB00138.1; JOINED.
 DR EMBL; L42459; AAB00138.1; JOINED.
 DR EMBL; L42460; AAB00138.1; JOINED.
 DR EMBL; L42461; AAB00138.1; JOINED.
 DR EMBL; AJ009862; CAA08900.1; -.
 DR PIR; A01396; WFMS2.
 DR HSSP; P01137; 1KLA.
 DR MGD; MGI:98725; Tgfb1.
 DR GO; GO:0005578; C:extracellular matrix; IDA.
 DR GO; GO:0006954; P:inflammatory response; IMP.
 DR GO; GO:0007515; P:lymph gland development; IMP.

DR GO; GO:0008220; P:necrosis; IMP.
 DR GO; GO:0016202; P:regulation of myogenesis; IDA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 278
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 390 AA; 44310 MW; 4381A51B711D689E CRC64;

Query Match 90.7%; Score 621; DB 1; Length 390;
 Best Local Similarity 92.5%; Pred. No. 5e-59;
 Matches 111; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330

 Qy 61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
 |||||:|||||
 Db 331 SLDTQYSKVLALYNQHNPGASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 8

TGF1_RAT

ID TGF1_RAT STANDARD; PRT; 390 AA.
 AC P17246;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).
 GN TGFB1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Heart;

```

RX  MEDLINE=90272425; PubMed=2349108;
RA  Qian S.W., Kondaiah P., Roberts A.B., Sporn M.B.;
RT  "cDNA cloning by PCR of rat transforming growth factor beta-1.";
RL  Nucleic Acids Res. 18:3059-3059(1990).
CC  -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC      PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC      TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC      HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC      ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC      A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC  -!- SUBUNIT: Homodimer; disulfide-linked.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: Belongs to the TGF-beta family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X52498; CAA36741.1; -.
DR  PIR; S10219; S10219.
DR  HSSP; P01137; 1KLA.
DR  InterPro; IPR002400; GF_cysknot.
DR  InterPro; IPR003911; TGF_TGFb.
DR  InterPro; IPR001839; TGFb.
DR  InterPro; IPR001111; TGFb_N.
DR  Pfam; PF00019; TGF-beta; 1.
DR  Pfam; PF00688; TGFb_propeptide; 1.
DR  PRINTS; PR00438; GFCYSKNOT.
DR  PRINTS; PR01423; TGFbBETA.
DR  ProDom; PD000357; TGFb; 1.
DR  SMART; SM00204; TGFb; 1.
DR  PROSITE; PS00250; TGF_BETA_1; 1.
KW  Growth factor; Mitogen; Glycoprotein; Signal.
FT  SIGNAL          1      23      POTENTIAL.
FT  PROPEP          24      278
FT  CHAIN           279      390      TRANSFORMING GROWTH FACTOR BETA 1.
FT  DISULFID        285      294      BY SIMILARITY.
FT  DISULFID        293      356      BY SIMILARITY.
FT  DISULFID        322      387      BY SIMILARITY.
FT  DISULFID        326      389      BY SIMILARITY.
FT  DISULFID        355      355      INTERCHAIN (BY SIMILARITY).
FT  CARBOHYD        82      82      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD       136     136      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD       176     176      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  SITE           244     246      CELL ATTACHMENT SITE (POTENTIAL).
SQ  SEQUENCE       390 AA;  44329 MW;  5E21108ED50D853C CRC64;

```

```

Query Match          90.7%;  Score 621;  DB 1;  Length 390;
Best Local Similarity 92.5%;  Pred. No. 5e-59;
Matches 111;  Conservative 1;  Mismatches 0;  Indels 8;  Gaps 1;

```

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      |||||  |||||

```

```

Db          279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330
QY          61 SLDTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
           |||||:|||||
Db          331 SLDTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

```

RESULT 9

TGF1_HORSE

```

ID  TGF1_HORSE      STANDARD;          PRT;    390 AA.
AC  O19011;
DT  15-JUL-1998 (Rel. 36, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Transforming growth factor beta 1 precursor (TGF-beta 1).
GN  TGFB1.
OS  Equus caballus (Horse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX  NCBI_TaxID=9796;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Lymph node;
RX  MEDLINE=98185507; PubMed=9524819;
RA  Penha-Goncalves M.N., Onions D.E., Nicolson L.;
RT  "Cloning and sequencing of equine transforming growth factor-beta 1
RT  (TGF beta-1) cDNA.";
RL  DNA Seq. 7:375-378(1997).
CC  -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC  PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC  TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC  HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC  ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC  A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC  -!- SUBUNIT: Homodimer; disulfide-linked.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- SIMILARITY: Belongs to the TGF-beta family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X99438; CAA67801.1; -.
DR  HSSP; P01137; 1KLA.
DR  InterPro; IPR002400; GF_cysknot.
DR  InterPro; IPR003911; TGF_TGFB.
DR  InterPro; IPR001839; TGFB.
DR  InterPro; IPR001111; TGFB_N.
DR  Pfam; PF00019; TGF-beta; 1.
DR  Pfam; PF00688; TGFB_propeptide; 1.
DR  PRINTS; PR00438; GFCYSKNOT.
DR  PRINTS; PR01423; TGFBETA.
DR  ProDom; PD000357; TGFB; 1.

```

DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 278 BY SIMILARITY.
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 390 AA; 43974 MW; A86D715F44549691 CRC64;

Query Match 90.5%; Score 620; DB 1; Length 390;
 Best Local Similarity 92.5%; Pred. No. 6.3e-59;
 Matches 111; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 279 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330
 Qy 61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
 |||||
 Db 331 SLDTQYSKVLALYNQHNP GASAAPCCVPQVLEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 10

TGF1_CAVPO

ID TGF1_CAVPO STANDARD; PRT; 390 AA.
 AC Q9Z1Y6; Q9QZB3; Q9R148;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transforming growth factor beta 1 precursor (TGF-beta 1).
 GN TGFB1.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Hartley;
 RA Jeevan A., McMurray D.N., Yoshimura T.;
 RT "Guinea pig transforming growth factor-beta in peritoneal exudates
 after BCG vaccination."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE OF 265-382 FROM N.A.
 RX MEDLINE=99144670; PubMed=10025978;
 RA Scarozza A.M., Ramsingh A.I., Wicher V., Wicher K.;
 RT "Spontaneous cytokine gene expression in normal guinea pig blood and
 tissues."
 RL Cytokine 10:851-859(1998).
 RN [3]

RP SEQUENCE OF 279-371 FROM N.A.
 RC STRAIN=Hartley; TISSUE=Trachea;
 RA Morishima Y., Uchida Y., Nomura A., Ishii Y., Sakamoto T.,
 RA Sekizawa K.;
 RT "Guinea-pig transforming growth factor-beta expression in injured
 RT tracheal epithelium.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: MULTIFUNCTIONAL PEPTIDE THAT CONTROLS PROLIFERATION,
 CC DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL TYPES. MANY
 CC CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF THEM HAVE
 CC SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1 REGULATES THE
 CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES A
 CC POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF191297; AAF02780.1; -.
 DR EMBL; AF097509; AAC83807.1; -.
 DR EMBL; AF169347; AAD49347.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT PROPEP 25 278 POTENTIAL.
 FT CHAIN 279 390 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 285 294 BY SIMILARITY.
 FT DISULFID 293 356 BY SIMILARITY.
 FT DISULFID 322 387 BY SIMILARITY.
 FT DISULFID 326 389 BY SIMILARITY.
 FT DISULFID 355 355 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 244 246 CELL ATTACHMENT SITE (POTENTIAL).
 FT CONFLICT 279 279 G -> P (IN REF. 3).
 FT CONFLICT 286 286 F -> S (IN REF. 2).
 FT CONFLICT 309 309 K -> E (IN REF. 2).
 FT CONFLICT 322 322 C -> R (IN REF. 2).

FT CONFLICT 350 350 A -> G (IN REF. 2).
SQ SEQUENCE 390 AA; 44328 MW; 1539F849BA0C0FF1 CRC64;

Query Match 89.3%; Score 612; DB 1; Length 390;
Best Local Similarity 92.4%; Pred. No. 4.5e-58;
Matches 110; Conservative 0; Mismatches 1; Indels 8; Gaps 1;

```
QY      2 LDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWS 61
      |||||
Db      280 LDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWS 331

QY      62 LDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      |||||
Db      332 LDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKAKVEQLSNMIVRSCKCS 390
```

RESULT 11

TGF1 CHICK

ID TGF1_CHICK STANDARD; PRT; 373 AA.
AC P09531;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transforming growth factor beta 1 precursor (TGF-beta 1) (TGF-beta 4)
DE (Fragment).
GN TGFB1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=89112198; PubMed=2464131;
RA Jakowlew S.B., Dillard P.J., Sporn M.B., Roberts A.B.;
RT "Complementary deoxyribonucleic acid cloning of a messenger
RT ribonucleic acid encoding transforming growth factor beta 4 from
RT chicken embryo chondrocytes."
RL Mol. Endocrinol. 2:1186-1195(1988).
RN [2]
RP REVISIONS.
RX MEDLINE=92357039; PubMed=1353860;
RA Burt D.W., Jakowlew S.B.;
RT "Correction: a new interpretation of a chicken transforming growth
RT factor-beta 4 complementary DNA."
RL Mol. Endocrinol. 6:989-992(1992).
CC -!- FUNCTION: TGF-BETA IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROL
CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA AND ESSENTIALLY ALL OF THEM
CC HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA REGULATES THE
CC ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND DETERMINES
CC A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the TGF-beta family.
CC -----

OS *Xenopus laevis* (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; *Xenopus*.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90110090; PubMed=2295601;
 RA Kondaiah P., Sands M.J., Smith J.M., Fields A., Roberts A.B.,
 RA Sporn M.B., Melton D.A.;
 RT "Identification of a novel transforming growth factor-beta (TGF-beta
 RT 5) mRNA in *Xenopus laevis*.";
 RL J. Biol. Chem. 265:1089-1093(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Vempati U.D., Kondaiah P.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: IMPORTANT ROLE IN CERTAIN ASPECTS OF DIFFERENTIATION
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J05180; AAA49968.1; -.
 DR EMBL; AF009335; AAB64441.1; -.
 DR EMBL; AF009331; AAB64441.1; JOINED.
 DR EMBL; AF009332; AAB64441.1; JOINED.
 DR EMBL; AF009333; AAB64441.1; JOINED.
 DR EMBL; AF009334; AAB64441.1; JOINED.
 DR PIR; A34929; B61036.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFbBETA.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 270
 FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 277 286 BY SIMILARITY.
 FT DISULFID 285 348 BY SIMILARITY.
 FT DISULFID 314 379 BY SIMILARITY.
 FT DISULFID 318 381 BY SIMILARITY.

FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 234 236 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 382 AA; 44200 MW; 1034621C917AAE15 CRC64;

Query Match 71.7%; Score 491; DB 1; Length 382;
 Best Local Similarity 73.0%; Pred. No. 3.9e-45;
 Matches 84; Conservative 10; Mismatches 13; Indels 8; Gaps 1;

Qy 6 YCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQ 65
 ||| : ||||: |||:||||||| ||: || |||||: |||
 Db 276 YCFGNG-----PNCCVKPLYINFRKDLGWKWIHEPKGYEANYCLGNCPYIWSMDTQ 327
 Qy 66 YSKVLALYNQHNPASAAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
 |||||: ||||: |||| : |||| |||||: |||| |||||: |||| ||
 Db 328 YSKVLSLYNQNNPGASISPCCVDPVLEPLPIIYVVRTAKVEQLSNMIVRSNCNS 382

RESULT 13

TGF3_MOUSE

ID TGF3_MOUSE STANDARD; PRT; 410 AA.
 AC P17125;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transforming growth factor beta 3 precursor (TGF-beta 3).
 GN TGFB3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90190650; PubMed=2628730;
 RA Miller D.A., Lee A., Matsui Y., Chen E.Y., Moses H.L., Derynck R.;
 RT "Complementary DNA cloning of the murine transforming growth
 RT factor-beta 3 (TGF beta 3) precursor and the comparative expression
 RT of TGF beta 3 and TGF beta 1 messenger RNA in murine embryos and
 RT adult tissues."
 RL Mol. Endocrinol. 3:1926-1934(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91000714; PubMed=2206556;
 RA Denhez F., Lafyatis R., Kondaiah P., Roberts A.B., Sporn M.B.;
 RT "Cloning by polymerase chain reaction of a new mouse TGF-beta,
 RT mTGF-beta 3."
 RL Growth Factors 3:139-146(1990).
 CC -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M32745; AAA40422.1; -.
DR PIR; A41397; A41397.
DR HSSP; P10600; 1TGJ.
DR MGD; MGI:98727; Tgfb3.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR003911; TGF_TGFB.
DR InterPro; IPR001839; TGFB.
DR InterPro; IPR001111; TGFB_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFB_propeptide; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRINTS; PR01423; TGFBETA.
DR ProDom; PD000357; TGFB; 1.
DR SMART; SM00204; TGFB; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Mitogen; Glycoprotein; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 298 POTENTIAL.
FT CHAIN 299 410 TRANSFORMING GROWTH FACTOR BETA 3.
FT DISULFID 305 314 BY SIMILARITY.
FT DISULFID 313 376 BY SIMILARITY.
FT DISULFID 342 407 BY SIMILARITY.
FT DISULFID 346 409 BY SIMILARITY.
FT DISULFID 375 375 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 259 261 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 410 AA; 46884 MW; 250F7048CA432BD6 CRC64;

Query Match 70.9%; Score 486; DB 1; Length 410;
Best Local Similarity 71.7%; Pred. No. 1.5e-44;
Matches 86; Conservative 11; Mismatches 15; Indels 8; Gaps 1;

QY 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
| | | | | : | : | | | | | | | : | | | | : | | | | :
Db 299 ALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKVVHEPKGYANFCSGPCPYLR 350

QY 61 SLDTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
| | : | | | | | | | : | | | | | | | : | | | | : | | | |
Db 351 SADTTHSTVLGLYNTLNPEASAPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS 410

RESULT 14

TGF3_HUMAN

ID TGF3_HUMAN STANDARD; PRT; 412 AA.
AC P10600;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transforming growth factor beta 3 precursor (TGF-beta 3).
GN TGFB3.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88263019; PubMed=3164476;
 RA ten Dijke P., Hansen P., Iwata K., Pieler C., Foulkes J.G.;
 RT "Identification of another member of the transforming growth factor
 RT type beta gene family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:4715-4719(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=89091120; PubMed=3208746;
 RA Derynck R., Lindquist P.B., Lee A., Wen D., Tamm J., Graycar J.L.,
 RA Rhee L., Mason A.J., Miller D.A., Coffey R.J., Moses H.L., Chen E.Y.;
 RT "A new type of transforming growth factor-beta, TGF-beta 3";
 RL EMBO J. 7:3737-3743(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Madan A., Rowen L., Qin S., Dickhoff R., Shaffer T., James R.,
 RA Abbasi N., Loretz C., Madan A., Dors M., Dahl T., Hall J., Lasky S.,
 RA Hood L.;
 RT "Complete genomic sequence of human transforming growth factor-beta
 RT 3.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
 RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
 RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 301-412.
 RX MEDLINE=96416253; PubMed=8819159;
 RA Mittl P.R., Priestle J.P., Cox D.A., McMaster G., Cerletti N.,
 RA Grutter M.G.;
 RT "The crystal structure of TGF-beta 3 and comparison to TGF-beta 2:
 RT implications for receptor binding.";
 RL Protein Sci. 5:1261-1271(1996).
 CC -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.
 CC -!- SUBUNIT: Homodimer; disulfide-linked.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the TGF-beta family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J03241; AAA61161.1; -.
 DR EMBL; X14149; CAA32362.1; -.
 DR EMBL; X14885; CAA33024.1; ALT_INIT.

DR EMBL; X14886; CAA33024.1; JOINED.
 DR EMBL; X14887; CAA33024.1; JOINED.
 DR EMBL; X14888; CAA33024.1; JOINED.
 DR EMBL; X14889; CAA33024.1; JOINED.
 DR EMBL; X14890; CAA33024.1; JOINED.
 DR EMBL; X14891; CAA33024.1; JOINED.
 DR EMBL; AF107885; AAC79727.1; -.
 DR EMBL; AY140241; AAM96819.1; -.
 DR PIR; A36169; A36169.
 DR PDB; 1TGJ; 11-JAN-97.
 DR PDB; 1TGK; 12-MAR-97.
 DR PDB; 1KTZ; 27-FEB-02.
 DR Genew; HGNC:11769; TGFB3.
 DR MIM; 190230; -.
 DR GO; GO:0005160; F:transforming growth factor-beta receptor li. . .; TAS.
 DR GO; GO:0007267; P:cell-cell signaling; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR003911; TGF_TGFB.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal; 3D-structure.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 300
 FT CHAIN 301 412 TRANSFORMING GROWTH FACTOR BETA 3.
 FT DISULFID 307 316
 FT DISULFID 315 378
 FT DISULFID 344 409
 FT DISULFID 348 411
 FT DISULFID 377 377 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 261 263 CELL ATTACHMENT SITE (POTENTIAL).
 FT STRAND 303 303
 FT HELIX 304 309
 FT STRAND 316 318
 FT STRAND 321 323
 FT HELIX 324 328
 FT TURN 331 332
 FT STRAND 333 335
 FT STRAND 338 340
 FT STRAND 343 345
 FT TURN 350 353
 FT STRAND 354 354
 FT HELIX 357 368
 FT TURN 370 371
 FT STRAND 378 380
 FT STRAND 383 392

FT	TURN	393	394	
FT	STRAND	395	406	
FT	STRAND	408	412	
SQ	SEQUENCE	412 AA;	47328 MW;	3CAD3548D3AEA178 CRC64;

Query Match 70.9%; Score 486; DB 1; Length 412;
Best Local Similarity 71.7%; Pred. No. 1.5e-44;
Matches 86; Conservative 11; Mismatches 15; Indels 8; Gaps 1;

Qy	1	ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW	60
		: : : : : :	
Db	301	ALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKWVHEPKGYANFCSGPCPYLR	352
Qy	61	SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS	120
		: : : : :	
Db	353	SADTTHTSTVLGLYNTLNPEASAPCCVPODLEPLTILYYVGRTPKVEQLSNMIVVKSCKCS	412

RESULT 15

```

TGF3_RAT
ID   TGF3_RAT          STANDARD;          PRT;   412 AA.
AC   Q07258;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Transforming growth factor beta 3 precursor (TGF-beta 3).
GN   TGFB3 OR TGF-B3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Lung;
RX   MEDLINE=95155340; PubMed=7852342;
RA   Wang J., Kuliszewski M., Yee W., Sedlackova L., Xu J., Tseu I.,
RA   Post M.;
RT   "Cloning and expression of glucocorticoid-induced genes in fetal rat
RT   lung fibroblasts. Transforming growth factor-beta 3.";
RL   J. Biol. Chem. 270:2722-2728(1995).
RN   [2]
RP   SEQUENCE OF 159-213 FROM N.A.
RX   MEDLINE=93286190; PubMed=8509457;
RA   McKinnon R.D., Piras G., Ida J., Dubois-Dalq M.;
RT   "A role for TGF-beta in oligodendrocyte differentiation.";
RL   J. Cell Biol. 121:1397-1407(1993).
CC   -!- FUNCTION: INVOLVED IN EMBRYOGENESIS AND CELL DIFFERENTIATION.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the TGF-beta family.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce

```

CC or send an email to license@isb-sib.ch).

```

DR      EMBL; U03491; AAA67915.1; -.
DR      EMBL; X71903; CAA50722.1; -.
DR      PIR; A55706; A55706.
DR      HSSP; P10600; 1TGJ.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR003911; TGF_TGFB.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFbBETA.
DR      ProDom; PD000357; TGFb; 1.
DR      SMART; SM00204; TGFB; 1.
DR      PROSITE; PS00250; TGF_BETA_1; 1.
KW      Growth factor; Mitogen; Glycoprotein; Signal.
FT      SIGNAL          1          23          POTENTIAL.
FT      PROPEP          24         300          POTENTIAL.
FT      CHAIN           301         412          TRANSFORMING GRO
FT      DISULFID        307         316          BY SIMILARITY.
FT      DISULFID        315         378          BY SIMILARITY.
FT      DISULFID        344         409          BY SIMILARITY.
FT      DISULFID        348         411          BY SIMILARITY.
FT      DISULFID        377         377          INTERCHAIN (BY S
FT      CARBOHYD        74          74          N-LINKED (GLCNAC
FT      CARBOHYD       135         135          N-LINKED (GLCNAC
FT      CARBOHYD       142         142          N-LINKED (GLCNAC
FT      SITE            261         263          CELL ATTACHMENT
SQ      SEQUENCE        412 AA;  47116 MW;  24FD7D899090AA

```

Qy	1	ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW	60
		: : : : :	
Db	301	ALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKWVHEPKGYANFCSGPCPYLR	352
Qy	61	SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYYVGRKPKVEQLSNMIVRSCKCS	120
		: : : :	
Db	353	SSDTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS	412

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

```
Run on:      October 28, 2003, 07:50:55 ; Search time 9.6208 Seconds
              (without alignments)
              3218.683 Million cell updates/sec
```

Title: US-10-017-372E-15
 Perfect score: 685
 Sequence: 1 ALDTNYCFSSTDYKDDDDKE.....GRKPKVEQLSNMIVRSCKCS 120

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : SPTREMBL_23:*
 1: sp_archaea:*
 2: sp_bacteria:*
 3: sp_fungi:*
 4: sp_human:*
 5: sp_invertebrate:*
 6: sp_mammal:*
 7: sp_mhc:*
 8: sp_organelle:*
 9: sp_phage:*
 10: sp_plant:*
 11: sp_rodent:*
 12: sp_virus:*
 13: sp Vertebrate:*
 14: sp_unclassified:*
 15: sp_rvirus:*
 16: sp_bacteriap:*
 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	624	91.1	112	6	O02730	O02730 oryctolagus	
2	621	90.7	368	11	Q8R4D9	Q8r4d9 sigmodon hi	
3	612	89.3	390	6	Q9TUM8	Q9tum8 equus cabal	
4	609	88.9	130	11	Q08714	Q08714 mesocricetu	
5	606	88.5	124	6	Q95N80	Q95n80 canis famil	
6	540	78.8	101	11	Q9R184	Q9r184 meriones un	
7	486	70.9	362	11	Q99K17	Q99k17 mus musculu	
8	486	70.9	412	11	Q91YU7	Q91yu7 mus musculu	
9	478	69.8	382	13	O93449	O93449 oncorhynchu	
10	466	68.0	255	11	Q921T1	Q921t1 mus musculu	

11	466	68.0	414	11	Q91VP5	Q91vp5 mus musculu
12	464.5	67.8	200	13	Q90YF1	Q90yf1 pleuronecte
13	463.5	67.7	379	13	Q8JHF5	Q8jhf5 sparus aura
14	462.5	67.5	382	13	Q9PWA9	Q9pwa9 morone chry
15	459.5	67.1	379	13	Q8AXK8	Q8axk8 sparus aura
16	454	66.3	224	11	Q8CDZ9	Q8cdz9 mus musculu
17	436	63.6	376	13	Q9PTQ2	Q9ptq2 cyprinus ca
18	424	61.9	361	13	Q98854	Q98854 cyprinus ca
19	420	61.3	399	11	Q9ERB7	Q9erb7 mesocricetu
20	413	60.3	88	13	Q90YF5	Q90yf5 pleuronecte
21	402	58.7	88	13	Q90YF7	Q90yf7 oncorhynchu
22	397	58.0	88	13	Q90ZE7	Q90ze7 acipenser b
23	393	57.4	87	13	Q42306	Q42306 carassius a
24	382	55.8	91	6	Q9MYZ1	Q9myz1 capra hircu
25	373	54.5	86	6	Q28241	Q28241 cervus elap
26	358	52.3	81	6	Q9N1S3	Q9nls3 capreolus c
27	353	51.5	179	13	Q90YF2	Q90yf2 pleuronecte
28	317	46.3	77	13	Q90YF8	Q90yf8 oncorhynchu
29	301	43.9	62	13	Q90ZJ7	Q90zj7 anguilla an
30	300	43.8	62	13	Q90YF4	Q90yf4 pleuronecte
31	289	42.2	50	6	Q28240	Q28240 cervus elap
32	287	41.9	62	13	Q9DEP5	Q9dep5 scophthalmu
33	283	41.3	62	13	Q90ZJ8	Q90zj8 anguilla an
34	283	41.3	62	13	Q90YF9	Q90yf9 oncorhynchu
35	273	39.9	62	13	Q90YF3	Q90yf3 pleuronecte
36	205.5	30.0	389	13	Q90YY0	Q90yy0 ictalurus p
37	205	29.9	426	4	Q9HBP0	Q9hbp0 homo sapien
38	204	29.8	424	13	Q98SP9	Q98sp9 meleagris g
39	203.5	29.7	385	13	Q90W05	Q90w05 sparus aura
40	202.5	29.6	374	13	Q8JFS0	Q8jfs0 brachydanio
41	201	29.3	119	6	Q95N79	Q95n79 ailuropoda
42	201	29.3	119	6	Q95KP2	Q95kp2 ursus malay
43	200.5	29.3	373	13	Q90ZD2	Q90zd2 oncorhynchu
44	200.5	29.3	373	13	Q90ZD1	Q90zd1 oncorhynchu
45	200.5	29.3	373	13	Q9DDI8	Q9ddi8 salmo salar

ALIGNMENTS

RESULT 1

O02730

ID O02730 PRELIMINARY; PRT; 112 AA.

AC O02730; O97501;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).

GN TGFB1 OR TGF-BETA-1.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RA Taylor T.K., James E.R., McGonigle S., Yoho E.R.;

RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.

RN [2]
 RP SEQUENCE OF 2-99 FROM N.A.
 RA Inoue K., Kawabe Y., Kodama T.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF
 CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1
 CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND
 CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AF000133; AAB53806.1; -.
 DR EMBL; AB020217; BAA36950.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein.
 FT NON_TER 1 1
 FT CHAIN 1 112 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 7 16 BY SIMILARITY.
 FT DISULFID 15 78 BY SIMILARITY.
 FT DISULFID 44 109 BY SIMILARITY.
 FT DISULFID 48 111 BY SIMILARITY.
 FT DISULFID 77 77 INTERCHAIN (BY SIMILARITY).
 FT CONFLICT 2 3 LD -> FS (IN REF. 2).
 FT CONFLICT 85 92 PLPIVYYV -> ATAHRTTTL (IN REF. 2).
 SQ SEQUENCE 112 AA; 12795 MW; 53C5B7D46355A6F3 CRC64;

Query Match 91.1%; Score 624; DB 6; Length 112;
 Best Local Similarity 93.3%; Pred. No. 4.6e-65;
 Matches 112; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 1 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 52
 Qy 61 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
 |||||
 Db 53 SLDTQYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 112

RESULT 2

Q8R4D9

ID Q8R4D9 PRELIMINARY; PRT; 368 AA.
 AC Q8R4D9;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transforming growth factor beta-1 protein (Fragment).
 GN TGFB1.
 OS Sigmodon hispidus (Hispid cotton rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
 OC Sigmodon.
 OX NCBI_TaxID=42415;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Blanco J.C., Pletneva L.M., Prince G.A.;
 RT "Cotton rat cytokines, chemokines, and interferons.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AF480858; AAL87199.1; -.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 368 AA; 41905 MW; A5C91207B0468B4A CRC64;

Query Match 90.7%; Score 621; DB 11; Length 368;
 Best Local Similarity 92.5%; Pred. No. 4e-64;
 Matches 111; Conservative 1; Mismatches 0; Indels 8; Gaps 1;

QY 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 ||||||||||| ||||||||||| ||||||||||| ||||||||||| |||||||||||
 Db 257 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 308
 QY 61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
 ||||||||||| ||||||||||| : ||||||||||| ||||||||||| ||||||||||| |||||||||||
 Db 309 SLDTQYSKVLALYNQHNP GASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 368

RESULT 3

Q9TUM8

ID Q9TUM8 PRELIMINARY; PRT; 390 AA.
 AC Q9TUM8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transforming growth factor beta 1.
 GN TGFB1.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nixon A.J., Brower-Toland B.T., Sandell L.J.;
 RT "Molecular cloning of equine transforming growth factor beta 1 reveals
 RT a divergent nucleotide structure that encodes a novel bioactive
 RT peptide among mammalian species.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

DR EMBL; AF175709; AAD49431.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR InterPro; IPR003911; TGF_TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFbBETA.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 SQ SEQUENCE 390 AA; 43860 MW; 220FE40DFCCA6016 CRC64;

Query Match 89.3%; Score 612; DB 6; Length 390;
 Best Local Similarity 91.7%; Pred. No. 4.8e-63;
 Matches 110; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

QY 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 ||||| ||| ||||||||||||||||||||||||||||||||||||
 Db 279 ALDTNYCSSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 330
 QY 61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS 120
 ||||||||||||||||||||||||||||||||||||
 Db 331 SLDTQYSKVLALYNQHNP GASAAPCCVPQVLEPLIVYYVGRKPKVEQLSNMIVRSCKCS 390

RESULT 4

Q08714

ID Q08714 PRELIMINARY; PRT; 130 AA.
 AC Q08714; O70331;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).
 GN TGFb1.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LVG (SYR);
 RX MEDLINE=93304479; PubMed=8317544;
 RA Wong D.T., Donoff R.B., Yang J., Song B.Z., Matossian K., Nagura N.,
 RA Elovic A., McBride J., Gallagher G., Todd R.;
 RT "Sequential expression of transforming growth factors alpha and beta 1
 RT by eosinophils during cutaneous wound healing in the hamster.";
 RL Am. J. Pathol. 143:130-142(1993).
 RN [2]
 RP SEQUENCE OF 26-115 FROM N.A.
 RC STRAIN=SYRIAN; TISSUE=SPLEEN;
 RX MEDLINE=98234044; PubMed=9573100;
 RA Melby P.C., Tryon V.V., Chandrasekar B., Freeman G.L.;
 RT "Cloning of Syrian hamster (Mesocricetus auratus) cytokine cDNAs and

RT analysis of cytokine mRNA expression in experimental visceral
 RT leishmaniasis.";
 RL Infect. Immun. 66:2135-2142(1998).
 CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF
 CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1
 CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND
 CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
 CC -!- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.
 DR EMBL; X60296; CAA42838.1; -.
 DR EMBL; AF046214; AAC40099.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR001839; TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein.
 FT NON_TER 1 1
 FT PROPEP <1 18
 FT CHAIN 19 130 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 25 34 BY SIMILARITY.
 FT DISULFID 33 96 BY SIMILARITY.
 FT DISULFID 66 129 BY SIMILARITY.
 FT DISULFID 95 95 INTERCHAIN (BY SIMILARITY).
 FT CONFLICT 93 93 G -> S (IN REF. 2).
 SQ SEQUENCE 130 AA; 14997 MW; 8B41DD6CF39CCA77 CRC64;

Query Match 88.9%; Score 609; DB 11; Length 130;
 Best Local Similarity 91.7%; Pred. No. 3.1e-63;
 Matches 110; Conservative 0; Mismatches 2; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 19 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 70
 Qy 61 SLDTQYSKVLALYNQHNPASAGPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
 |||||
 Db 71 SLDTQYSKVLALYNQHNPASAGPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSYKCS 130

RESULT 5

Q95N80

ID Q95N80 PRELIMINARY; PRT; 124 AA.
 AC Q95N80;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transforming growth factor beta 1 (Fragment).
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Fonfara S., Groene A., Baumgaertner W.;
 RT "Sequence of canine transforming growth factor beta 1 mRNA in DH82-
 RT cells.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AF349538; AAK54072.1; -.
 DR InterPro; IPR001839; TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 FT NON_TER 1 1
 FT NON_TER 124 124
 SQ SEQUENCE 124 AA; 14329 MW; 21D185218E5556DB CRC64;

Query Match 88.5%; Score 606; DB 6; Length 124;
 Best Local Similarity 93.2%; Pred. No. 6.6e-63;
 Matches 109; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 |||||
 Db 16 ALDTNYCFSST-----EKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 67
 Qy 61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSC 117
 |||||
 Db 68 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSC 124

RESULT 6

Q9R184

ID Q9R184 PRELIMINARY; PRT; 101 AA.
 AC Q9R184;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Transforming growth factor beta 1 (TGF-beta 1) (Fragment).
 GN TGFB1 OR TGF-BETA.
 OS Meriones unguiculatus (Mongolian jird).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Meriones.
 OX NCBI_TaxID=10047;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LYMPH NODE;
 RA Rao U.R., Klei T.R.;
 RT "cDNA cloning of gerbil transforming growth factor-beta by PCR.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: TGF-BETA 1 IS A MULTIFUNCTIONAL PEPTIDE THAT CONTROLS
 CC PROLIFERATION, DIFFERENTIATION, AND OTHER FUNCTIONS IN MANY CELL
 CC TYPES. MANY CELLS SYNTHESIZE TGF-BETA 1 AND ESSENTIALLY ALL OF
 CC THEM HAVE SPECIFIC RECEPTORS FOR THIS PEPTIDE. TGF-BETA 1
 CC REGULATES THE ACTIONS OF MANY OTHER PEPTIDE GROWTH FACTORS AND
 CC DETERMINES A POSITIVE OR NEGATIVE DIRECTION OF THEIR EFFECTS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AF161218; AAD45726.1; -.

DR HSSP; P01137; 1KLA.
 DR InterPro; IPR001839; TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein.
 FT NON_TER 1 1
 FT CHAIN <1 >101 TRANSFORMING GROWTH FACTOR BETA 1.
 FT DISULFID 1 10 BY SIMILARITY.
 FT DISULFID 9 72 BY SIMILARITY.
 FT DISULFID 71 71 INTERCHAIN (BY SIMILARITY).
 FT NON_TER 101 101
 SQ SEQUENCE 101 AA; 11724 MW; ABF1CFDA264AEFED CRC64;

 Query Match 78.8%; Score 540; DB 11; Length 101;
 Best Local Similarity 89.0%; Pred. No. 2.7e-55;
 Matches 97; Conservative 1; Mismatches 3; Indels 8; Gaps 1;

QY 7 CFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQY 66
 ||||| ||||||||| ||||||||||||||||||||| |||||
 Db 1 CFSST-----EKNCCVRQLYRDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLHTQY 52

 QY 67 SKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVR 115
 |||||||||||||:||||||||||||||||||||||| ||
 Db 53 SKVLALYNQHNPGASASPCCVPQALEPLPIVYYVGRKPKVEQLSNMFVR 101

RESULT 7

Q99K17

ID Q99K17 PRELIMINARY; PRT; 362 AA.
 AC Q99K17;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to transforming growth factor, beta 3 (Fragment).
 GN TGFB3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; BC005513; AAH05513.1; -.
 DR HSSP; P10600; 1TGJ.
 DR MGD; MGI:98727; Tgfb3.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR InterPro; IPR003911; TGF_TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.

DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFB; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
FT NON_TER 1 1
SQ SEQUENCE 362 AA; 41486 MW; 0808E46180FDAE70 CRC64;

Query Match 70.9%; Score 486; DB 11; Length 362;
Best Local Similarity 71.7%; Pred. No. 2.3e-48;
Matches 86; Conservative 11; Mismatches 15; Indels 8; Gaps 1;

QY 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 251 ALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKVVHEPKGYANFCSGPCPYLR 302

QY 61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
| | : | | | | | | : ||| : ||| : ||| : ||| : ||| :
Db 303 SADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKCS 362

RESULT 8

Q91YU7

ID Q91YU7 PRELIMINARY; PRT; 412 AA.
AC Q91YU7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Transforming growth factor, beta 3.
GN TGFB3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; BC014690; AAH14690.1; -.
DR MGD; MGI:98727; Tgfb3.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR InterPro; IPR003911; TGF_TGFb.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRINTS; PR01423; TGFBETA.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFB; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
SQ SEQUENCE 412 AA; 47144 MW; F3EB65D046DF32AD CRC64;

Query Match 70.9%; Score 486; DB 11; Length 412;
Best Local Similarity 71.7%; Pred. No. 2.6e-48;
Matches 86; Conservative 11; Mismatches 15; Indels 8; Gaps 1;

QY 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60

```

      ||||| :      |:||||| |||||:|||||:|||||:||||| |||||:
Db      301 ALDTNYCFRNL-----EENCCVRPLYIDFRQDLGWKVVHEPKGYANFCSGPCPYLR 352

Qy      61 SLDTQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      ||| :| || ||| || |||:||||| ||||| :||| ||||| |||||:|||:|||
Db      353 SADTTHSTVLGLYNTLNPEASASPCCVPQDLEPLTILYYVGRTPKVEQLSNMVVKSCCKS 412

```

RESULT 9

O93449

```

ID      O93449      PRELIMINARY;      PRT;      382 AA.
AC      O93449; Q91217;
DT      01-NOV-1998 (TrEMBLrel. 08, Created)
DT      01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT      01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE      Transforming growth factor beta precursor.
GN      TGF-BETA OR TGF.
OS      Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC      Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX      NCBI_TaxID=8022;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=LEUKOCYTE;
RX      MEDLINE=99242020; PubMed=10227481;
RA      Daniels G.D., Secombes C.J.;
RT      "Genomic organisation of rainbow trout, Oncorhynchus mykiss TGF-
RT      BETA.";
RL      Dev. Comp. Immunol. 23:139-147(1999).
RN      [2]
RP      SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC      TISSUE=LEUKOCYTE;
RX      MEDLINE=98390168; PubMed=9722928;
RA      Hardie L.J., Laing K.J., Daniels G.D., Grabowski P.S., Cunningham C.,
RA      Secombes C.J.;
RT      "Isolation of the first piscine transforming growth factor beta gene:
RT      analysis reveals tissue specific expression and a potential regulatory
RT      sequence in rainbow trout (Oncorhynchus mykiss).";
RL      Cytokine 10:555-563(1998).
CC      -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE
CC      RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.
CC      -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC      -!- TISSUE SPECIFICITY: EXPRESSED IN BLOOD LEUKOCYTES, KIDNEY
CC      MACROPHAGES, BRAIN, GILL AND SPLEEN BUT NOT IN LIVER.
CC      -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR      EMBL; AJ007836; CAA07707.1; -.
DR      EMBL; X99303; CAA67685.1; -.
DR      HSSP; P01137; 1KLA.
DR      InterPro; IPR002400; GF_cysknot.
DR      InterPro; IPR001839; TGFb.
DR      InterPro; IPR001111; TGFb_N.
DR      InterPro; IPR003911; TGF_TGFb.
DR      Pfam; PF00019; TGF-beta; 1.
DR      Pfam; PF00688; TGFb_propeptide; 1.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PRINTS; PR01423; TGFbeta.

```

DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 270
 FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA.
 FT DISULFID 278 286 BY SIMILARITY.
 FT DISULFID 285 348 BY SIMILARITY.
 FT DISULFID 314 379 BY SIMILARITY.
 FT DISULFID 318 381 BY SIMILARITY.
 FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 237 237 N -> D (IN REF. 2).
 FT CONFLICT 345 345 Q -> H (IN REF. 2).
 FT CONFLICT 371 372 LS -> VP (IN REF. 2).
 FT CONFLICT 377 377 K -> M (IN REF. 2).
 SQ SEQUENCE 382 AA; 44136 MW; 93BD4D3540084B92 CRC64;

Query Match 69.8%; Score 478; DB 13; Length 382;
 Best Local Similarity 76.7%; Pred. No. 2.1e-47;
 Matches 79; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 18 DKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSKVLALYNQHN 77
 || ::|||:||||| || ||:|:| |||: : :|:| || ||
 Db 280 DKSESCCVRKLYIDFRKDLGWKWIHEPTGYFANYCIGPCTYIWNTENKYSQVLALYKHHN 339
 QY 78 PGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
 |||| | ||||| :|||: ||||| |||:|:|
 Db 340 PGASAQPCVPQVLEPLPIIYVGRQHKVEQLSNMIVKSCRC 382

RESULT 10

Q921T1

ID Q921T1 PRELIMINARY; PRT; 255 AA.
 AC Q921T1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to transforming growth factor, beta 2.
 GN TGFB2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; BC011055; AAH11055.1; -.
 DR MGD; MGI:98726; Tgfb2.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.

DR InterPro; IPR003911; TGF_TGFB.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 SQ SEQUENCE 255 AA; 29087 MW; 94540017F3C5C219 CRC64;

Query Match 68.0%; Score 466; DB 11; Length 255;
 Best Local Similarity 66.7%; Pred. No. 3.3e-46;
 Matches 80; Conservative 15; Mismatches 17; Indels 8; Gaps 1;

Qy 1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
 ||| ||| : : |||:| |||||:|||||||:|||| | |||:
 Db 144 ALDAAYCFRNV-----QDNCCLRPLYIDFKRDLGWKWIHEPKGYNANFCAGACPYLW 195

Qy 61 SLDTQYSKVLALYNQHNPASAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
 | |||:| |||:| || |||:| ||| | ||| |:| |:| | |:| |||||:| |||||
 Db 196 SSDTQHTKVLSTYNTINPEASAPCCVSQDLEPLTILYYIGNTPKIEQLSNMIVKSKCS 255

RESULT 11

Q91VP5

ID Q91VP5 PRELIMINARY; PRT; 414 AA.
 AC Q91VP5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to transforming growth factor, beta 2.
 GN TGFB2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; BC011170; AAH11170.1; -.
 DR MGD; MGI:98726; Tgfb2.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFB.
 DR InterPro; IPR001111; TGFB_N.
 DR InterPro; IPR003911; TGF_TGFB.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFB_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFBETA.
 DR ProDom; PD000357; TGFB; 1.
 DR SMART; SM00204; TGFB; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 SQ SEQUENCE 414 AA; 47588 MW; DB37A7C38881F286 CRC64;

Query Match 68.0%; Score 466; DB 11; Length 414;
 Best Local Similarity 66.7%; Pred. No. 5.7e-46;
 Matches 80; Conservative 15; Mismatches 17; Indels 8; Gaps 1;

```

Qy      1 ALDTNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIW 60
      ||| ||| :      : |||:| |||||::| ||||| |||||: ||| | |||:|
Db      303 ALDAAAYCFRNV-----QDNCCLRLPLYIDFKRDLGWKWIHEPKGYNANFCAGACPYLW 354

Qy      61 SLDTQYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS 120
      | |||::|||:||| || |||:|||| | |||| |::| ||:|||||||:|||||
Db      355 SSDTQHTKVLSLYNTINPEASASPCCVSQDLEPLTILYYIGNTPKIEQLSNMIVKSKCKS 414
  
```

RESULT 12

Q90YF1

```

ID   Q90YF1      PRELIMINARY;          PRT;   200 AA.
AC   Q90YF1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Transforming growth factor beta 1 (Fragment).
GN   TGF-BETA1.
OS   Pleuronectes platessa (Plaice).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC   Pleuronectoidei; Pleuronectidae; Pleuronectes.
OX   NCBI_TaxID=8262;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Laing K.J.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20394651; PubMed=10938738;
RA   Laing K.J., Cunningham C., Secombes C.J.;
RT   "Genes for three different isoforms of transforming growth factor-beta
RT   are present in plaice (Pleuronectes platessa) DNA.";
RL   Fish and Shellfish Immunol. 10:261-271(2000).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AJ318933; CAC60268.1; -.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR001839; TGFb.
DR   Pfam; PF00019; TGF-beta; 1.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
FT   NON_TER      1      1
FT   NON_TER      200    200
SQ   SEQUENCE      200 AA;  22851 MW;  4876FEB6A263B4CC CRC64;
  
```

Query Match 67.8%; Score 464.5; DB 13; Length 200;
 Best Local Similarity 67.5%; Pred. No. 3.8e-46;
 Matches 79; Conservative 15; Mismatches 14; Indels 9; Gaps 1;

```

Qy      4 TNYCFSSTDYKDDDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD 63
  
```

```

Db          93 TDTCTAQTE-----TCCVRKLYIDFRKDLGWKWIHKPTGYHANYCMGSCTYIWNAE 143
Qy          64 TQYSKVLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
              :|||:||||| ||||| ||||| |||||:|||||: ||||| ||:|||||
Db          144 NKYSQILALYKHHNPGASAPCCVPQALEPLPILYYVGRQHKVEQLSNMSVKSCCKCS 200

```

RESULT 13

Q8JHF5

```

ID   Q8JHF5          PRELIMINARY;          PRT;    379 AA.
AC   Q8JHF5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Transforming growth factor beta 1.
OS   Sparus aurata (Gilthead sea bream).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percidae;
OC   Sparidae; Sparus.
OX   NCBI_TaxID=8175;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tafalla C., Aranguren R., Secombes C.J., Castrillo J.L., Novoa B.,
RA   Figueras A.;
RT   "Molecular characterization of sea bream (Sparus aurata) transforming
RT   growth factor beta1.";
RL   Fish and Shellfish Immunol. 0:0-0(2002).
CC   -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR   EMBL; AF510084; AAN03842.1; -.
DR   InterPro; IPR002400; GF_cysknot.
DR   InterPro; IPR001839; TGFb.
DR   InterPro; IPR001111; TGFb_N.
DR   InterPro; IPR003911; TGF_TGFb.
DR   Pfam; PF00019; TGF-beta; 1.
DR   Pfam; PF00688; TGFb_propeptide; 2.
DR   PRINTS; PR00438; GFCYSKNOT.
DR   PRINTS; PR01423; TGFBETA.
DR   ProDom; PD000357; TGFb; 1.
DR   SMART; SM00204; TGFB; 1.
DR   PROSITE; PS00250; TGF_BETA_1; 1.
SQ   SEQUENCE    379 AA;  43506 MW;  C0C9D3D2FCA29C0E CRC64;

```

```

Query Match          67.7%;  Score 463.5;  DB 13;  Length 379;
Best Local Similarity 71.4%;  Pred. No. 1e-45;
Matches    80;  Conservative 14;  Mismatches 17;  Indels    1;  Gaps    1;

```

```

Qy          10 STDYKDD-DDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK 68
              ||: ||   : : |||| ||||| ||||| |||||:| ||||:|:| | |||: : :||:
Db          268 STETKDTCTAQTTCCVRSLYIDFRKDLGWKWIHKPTRYHANYCMGSCTYIWNENKYSQ 327
Qy          69 VLALYNQHNPASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
              :|||| | ||||| ||||| |||||:|||||: ||||| |||||:|||||
Db          328 ILALYKHHNPGASAPCCVPQALEPLPILYYVGRQHKVEQLSNMIVKSCCKCS 379

```


RESULT 14

Q9PWA9

ID Q9PWA9 PRELIMINARY; PRT; 382 AA.
 AC Q9PWA9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transforming growth factor beta precursor.
 GN TGF-BETA.
 OS Morone chrysops x Morone saxatilis (white bass x striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OX NCBI_TaxID=45352;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=KIDNEY;
 RX MEDLINE=20394636; PubMed=10938723;
 RA Harms C.A., Kennedy-Stoskopf S., Horne W.A., Fuller F.J.,
 RA Tompkins W.A.F.;
 RT "Cloning and sequencing hybrid striped bass (Morone saxatilis x M.
 RT chrysops) transforming growth factor-beta (TGF-beta), and development
 RT of a reverse transcription quantitative competitive polymerase chain
 RT reaction (RT-qcPCR) assay to measure TGF-beta mRNA of teleost fish.";
 RL Fish Shellfish Immunol. 10:61-85(2000).
 CC -!- FUNCTION: IS LIKELY TO BE AN IMPORTANT CYTOKINE REGULATING IMMUNE
 CC RESPONSE. MAY ALSO HAVE A ROLE IN OTHER PHYSIOLOGICAL SYSTEMS.
 CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: HIGHER LEVELS FOUND IN MONONUCLEAR CELLS FROM
 CC PERIPHERAL BLOOD THAN IN SPLEEN OR ANTERIOR KIDNEY.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AF140363; AAD46997.1; -.
 DR HSSP; P01137; 1KLA.
 DR InterPro; IPR002400; GF_cysknot.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR InterPro; IPR003911; TGF_TGFb.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PRINTS; PR01423; TGFbBETA.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Growth factor; Mitogen; Glycoprotein; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT PROPEP ? 270
 FT CHAIN 271 382 TRANSFORMING GROWTH FACTOR BETA.
 FT DISULFID 278 286 BY SIMILARITY.
 FT DISULFID 285 348 BY SIMILARITY.
 FT DISULFID 314 379 BY SIMILARITY.
 FT DISULFID 318 381 BY SIMILARITY.
 FT DISULFID 347 347 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 235 237 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 382 AA; 43846 MW; 3124D8C34EA74D72 CRC64;

Query Match 67.5%; Score 462.5; DB 13; Length 382;
Best Local Similarity 71.4%; Pred. No. 1.3e-45;
Matches 80; Conservative 15; Mismatches 16; Indels 1; Gaps 1;

Qy 10 STDYKD-DDDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK 68
||: || : : |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : ||:
Db 271 STETKDVCTAQTETCCVRSlyIDFRKDLGWKWIHKPTGYNANYCMGSCTYIWNNAENKYSQ 330

Qy 69 VLALYNQHNPgASAApCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
: |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 ILALYKHHNPgASAApCCVPQALEPLPILYVGRQHKVEQLSNMIVKSKCS 382

RESULT 15

Q8AXK8

ID Q8AXK8 PRELIMINARY; PRT; 379 AA.
AC Q8AXK8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Transforming growth factor beta 1.
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RA Tafalla C., Novoa B., Aranguren R., Figueras A.;
RT "Molecular cloning and characterization of sea bream (Sparus aurata)
RT TGF beta 1.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF424703; AAN76665.1; -.
SQ SEQUENCE 379 AA; 43485 MW; A306EC387F6DBA7C CRC64;

Query Match 67.1%; Score 459.5; DB 13; Length 379;
Best Local Similarity 71.4%; Pred. No. 3e-45;
Matches 80; Conservative 13; Mismatches 18; Indels 1; Gaps 1;

Qy 10 STDYKDD-DDKEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDTQYSK 68
||: || : : |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| : : ||:
Db 268 STETKDTCTAQTETCCVRSlyIDFRKDLGWKWIHKPTRYHANYCMGSCTYIWNNAENKYSQ 327

Qy 69 VLALYNQHNPgASAApCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS 120
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 328 XLALYKHHNPgASAApCCVPQALEPLPILYVGRQHKVEQLSNMIVKSKCS 379

Search completed: October 28, 2003, 09:12:25
Job time : 9.6208 secs

Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

10489.161 Million cell updates/sec

Sequence: 1 tggtagcgagatggcgctt.....cgattaaagcgccgcgact 1349

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Maximum DB seq length: 2000000000

Listing first 45 summaries

```

1:  gb_ba:*
2:  gb_htg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

```

```

28:  em_un:*
29:  em_vi:*
30:  em_htg_hum:*
31:  em_htg_inv:*
32:  em_htg_other:*
33:  em_htg_mus:*
34:  em_htg_pln:*
35:  em_htg_rod:*
36:  em_htg_mam:*
37:  em_htg_vrt:*
38:  em_sy:*
39:  em_htgo_hum:*
40:  em_htgo_mus:*
41:  em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	1233.2	91.4	3206	4	PIGTGFB1A	M23703 Sus scrofa	
2	1223.8	90.7	1326	6	AX338213	AX338213 Sequence	
3	1213.8	90.0	1750	4	GGTGFB1	X12373 Porcine mRNA	
4	1143.6	84.8	1605	4	SSTGFB1	Y00111 Porcine mRNA	
5	1122.2	83.2	2221	4	AF461808	AF461808 Sus scrof	
6	1002.6	74.3	1369	4	DOGTGFB1A	L34956 Canine tran	
7	991.6	73.5	2527	6	E00973	E00973 cDNA encodi	
8	990.8	73.4	2537	6	A06669	A06669 Synthetic m	
9	984.8	73.0	1173	4	OATGFB1	X76916 O.aries mRNA	
10	984.4	73.0	1780	9	BC000125	BC000125 Homo sapi	
11	984.4	73.0	1780	9	BC001180	BC001180 Homo sapi	
12	983	72.9	1561	9	AGMTGFB	M16658 Simian tran	
13	980.2	72.7	1821	6	E03028	E03028 DNA encodin	
14	979.8	72.6	1746	9	BC022242	BC022242 Homo sapi	
15	978.6	72.5	1560	6	I06216	I06216 Sequence 2	
16	977.8	72.5	2745	9	HSTGFB1	X02812 Human mRNA	
17	976.6	72.4	1560	6	I08268	I08268 Sequence 2	
18	959.2	71.1	1569	6	I06221	I06221 Sequence 3	
19	954.2	70.7	1561	6	I08275	I08275 Sequence 3	
20	939.6	69.7	1173	9	BT007245	BT007245 Homo sapi	
21	939.6	69.7	1173	12	BT007866	BT007866 Synthetic	
22	930.4	69.0	1173	4	ECRGFB1	X99438 E.caballus	
23	930.2	69.0	1176	6	AX615127	AX615127 Sequence	
24	928.6	68.8	1176	6	AX481432	AX481432 Sequence	
25	928.6	68.8	1176	6	AX615128	AX615128 Sequence	
26	924	68.5	1187	4	AF175709	AF175709 Equus cab	
27	901.2	66.8	1597	10	AF191297	AF191297 Cavia por	
28	872.8	64.7	1641	10	BC013738	BC013738 Mus muscu	
29	859	63.7	1579	10	MUSTGFRNA	M13177 Mouse trans	
30	859	63.7	2094	10	MMU009862	AJ009862 Mus muscu	
31	839.4	62.2	1585	10	RNTGFB1	X52498 Rat mRNA fo	
32	827.6	61.3	1117	4	BOVTGFB	M36271 Bovine tran	
33	803	59.5	1125	10	AF480858	AF480858 Sigmodon	

34	682.2	50.6	1675	6	I03310	I03310 Sequence 1
35	654	48.5	1376	6	AX528533	AX528533 Sequence
36	652.4	48.4	1389	6	AX528619	AX528619 Sequence
37	591.4	43.8	1352	6	AX528535	AX528535 Sequence
38	586.6	43.5	1350	6	AX528615	AX528615 Sequence
39	555.6	41.2	699	6	I05434	I05434 Sequence 4
40	366.4	27.2	489	6	AX455100	AX455100 Sequence
41	357.6	26.5	650	6	AX336646	AX336646 Sequence
42	357.6	26.5	650	9	HUMTGFBFA	M38449 Homo sapien
43	357.6	26.5	862	6	I03312	I03312 Sequence 3
44	339.2	25.1	469	10	MATGFB1	X60296 M.auratus m
45	331.8	24.6	1256	5	CHKTGFB4	M31160 Gallus gall

ALIGNMENTS

RESULT 1

PIGTGFB1A

LOCUS PIGTGFB1A 3206 bp mRNA linear MAM 31-MAR-1995

DEFINITION Sus scrofa transforming growth factor beta-1 mRNA, complete cds.

ACCESSION M23703

VERSION M23703.1 GI:755044

KEYWORDS transforming growth factor-beta-1.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 3206)

AUTHORS Kondaiah,P., Van Obberghen-Schilling,E., Ludwig,R.L., Dhar,R., Sporn,M.B. and Roberts,A.B.

TITLE cDNA cloning of porcine transforming growth factor-beta 1 mRNAs. Evidence for alternate splicing and polyadenylation

JOURNAL J. Biol. Chem. 263 (34), 18313-18317 (1988)

MEDLINE 89054010

PUBMED 2461367

COMMENT On Apr 1, 1995 this sequence version replaced gi:341017.

Original source text: Sus scrofa (strain miniature swine) cDNA to mRNA.

FEATURES

source

Location/Qualifiers

1. .3206

/organism="Sus scrofa"

/mol_type="mRNA"

/strain="miniature swine"

/db_xref="taxon:9823"

/cell_type="peripheral blood lymphocyte"

gene

1. .3206

/gene="TGF-beta-1"

CDS

906. .2078

/gene="TGF-beta-1"

/codon_start=1

/product="transforming growth factor-beta-1"

/protein_id="AAA64616.1"

/db_xref="GI:755045"

/translation="MPPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK
RIEAIHQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD
YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL

```

KLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR
LSAHCSCDSKDNTHLHVEINGFNSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYVGRKPKVEQLSNMIVRSCKCS"
polyA_signal 3186. .3191
               /gene="TGF-beta-1"
polyA_site   3206
               /gene="TGF-beta-1"
BASE COUNT   645 a   1041 c   924 g   596 t
ORIGIN

```

```

Query Match          91.4%;  Score 1233.2;  DB 4;  Length 3206;
Best Local Similarity 96.1%;  Pred. No. 3.7e-208;
Matches 1290;  Conservative 0;  Mismatches 13;  Indels 39;  Gaps 1;

```

```

Qy      6  CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGGCTGCTGTGGCTGC  65
      ||  ||| |||||
Db     901  CCCCCATGCCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGGCTGCTGTGGCTGC  960

Qy     66  TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG  125
      |||||
Db    961  TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG  1020

Qy    126  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC  185
      |||||
Db   1021  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC  1080

Qy    186  TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG  245
      | |||||
Db   1081  TCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG  1140

Qy    246  CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGGAGCCCGAGC  305
      |||||
Db   1141  CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGGAGCCCGAGC  1200

Qy    306  CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC  365
      |||||
Db   1201  CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC  1260

Qy    366  AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG  425
      |||||
Db   1261  AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG  1320

Qy    426  AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA  485
      |||||
Db   1321  AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA  1380

Qy    486  GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT  545
      |||||
Db   1381  GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT  1440

Qy    546  GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCCTTTG  605
      |||||
Db   1441  GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCCTTTG  1500

Qy    606  ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC  665
      |||||

```

Db	1501	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	1560
Qy	666	TCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1561	TCAGTGCCCACTGTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	1620
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTTGC	785
Db	1621	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTTGC	1680
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1681	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGA-	1739
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1740	-----GCCCTGGATACCAACTACTGCT	1761
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1762	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	1821
Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1822	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1881
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1882	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1941
Qy	1086	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1145
Db	1942	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	2001
Qy	1146	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1205
Db	2002	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	2061
Qy	1206	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCCACAGCCCCGCCACCCGGCAGGCCCGGC	1265
Db	2062	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCCACAGCCCCGCCACCCGGCAGGCCCGGC	2121
Qy	1266	CCCACCCCCGCCCCCCTCACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCCACTTG	1325
Db	2122	CCCACCCCCGCCCCCCTCACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCCACTTG	2181
Qy	1326	GGATCGATTAAAGCGGCCGGA	1347
Db	2182	GGATCGATTAAAGGTGGAGAGA	2203

RESULT 2

AX338213

LOCUS	AX338213	1326 bp	DNA	linear	PAT 09-JAN-2002
DEFINITION	Sequence 1 from Patent WO0181404.				
ACCESSION	AX338213				
VERSION	AX338213.1 GI:18128750				

KEYWORDS .

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1

AUTHORS Strober,W., Nakamura,K., Kitani,A. and Fuss,I.J.

TITLE Inducible plasmid vector encoding tgf_g(b) and uses thereof

JOURNAL Patent: WO 0181404-A 1 01-NOV-2001;
THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)

FEATURES Location/Qualifiers

source 1..1326
/organism="Sus scrofa"
/mol_type="genomic DNA"
/db_xref="taxon:9823"

CDS 16..1188
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD20538.1"
/db_xref="GI:18128751"
/translation="MAPSGLRLLPLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK
RIEAIRGQILSKRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD
YYAKEVTRVLMVESGNIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNDSWRYLSNRLLAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR
LSAHSSSDSKDNTLHVEINGFNSGRRGLATIHGMRPFLLLMATPLERAQHLHSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"

BASE COUNT 263 a 438 c 392 g 233 t

ORIGIN

Query Match 90.7%; Score 1223.8; DB 6; Length 1326;
Best Local Similarity 96.4%; Pred. No. 1.9e-206;
Matches 1299; Conservative 0; Mismatches 7; Indels 41; Gaps 3;

Qy	1	TGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTG	60
Db	6	TGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTG	65
Qy	61	GCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGA	120
Db	66	GCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGA	125
Qy	121	CATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCT	180
Db	126	CATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCT	185
Qy	181	TCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGT	240
Db	186	TCGGCTCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCCGT	245
Qy	241	ACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCC	300
Db	246	ACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCC	305
Qy	301	CGAGCCAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGG	360
Db	306	CGAGCCAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGG	365

Qy	361	CAACCAAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACAC	420
Db	366	CAACCAAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACAC	425
Qy	421	GTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCT	480
Db	426	GTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCT	485
Qy	481	GCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGA	540
Db	486	GCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGA	545
Qy	541	TTCCTGGCGCTACCTCAGCAACCGGTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTC	600
Db	546	TTCCTGGCGCTACCTCAGCAACCGGTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTC	605
Qy	601	CTTTGATGTACACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTT	660
Db	606	CTTTGATGTACACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTT	665
Qy	661	TCGCCTCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCAGTGGAAATTAA	720
Db	666	TCGCCTCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCAGTGGAAATTAA	725
Qy	721	CGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTT	780
Db	726	CGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTT	785
Qy	781	CCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCG	840
Db	786	CCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCG	845
Qy	841	CCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTA	900
Db	846	CCGA-----GCCCTGGATACCAACTA	866
Qy	901	CTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAA	960
Db	867	CTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAA	926
Qy	961	GGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGG	1020
Db	927	GGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGG	986
Qy	1021	GCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAA	1080
Db	987	GCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAA	1046
Qy	1081	CCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACT	1140
Db	1047	CCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACT	1106
Qy	1141	GCCCATCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGT	1200
Db	1107	GCCCATCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGT	1166

```

Qy      1201 GCGTTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCCCACAGCCCCGCCCCACCCGGCAGGC 1260
          |||
Db      1167 GCGTTCCTGCAAGTGCAGCTGA-GCCCCGCCCCGCCCCACAGCCCCGCCCCACCCGGCAGGC 1225

Qy      1261 CCGGCCCCACCCCGCCCGCCTACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCC 1320
          |||
Db      1226 CCGGCCCCACCCCGCCCGCCTACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCC 1285

Qy      1321 ACTTGGGATCGATTAAAGCGGCCGCGA 1347
          ||
Db      1286 AC-TGGGATCGATTAAAGGTGGAGAGA 1311

```

RESULT 3

GGTGFB1

LOCUS GGTGFB1 1750 bp mRNA linear MAM 27-MAR-1996

DEFINITION Porcine mRNA for transforming growth factor-beta 1.

ACCESSION X12373

VERSION X12373.1 GI:63808

KEYWORDS transforming growth factor-beta 1.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 1750)

AUTHORS Jakowlew,S.B., Dillard,P.J., Sporn,M.B. and Roberts,A.B.

TITLE Nucleotide sequence of chicken transforming growth factor-beta 1
(TGF-beta 1)

JOURNAL Nucleic Acids Res. 16 (17), 8730 (1988)

MEDLINE 88335639

PUBMED 3166520

REFERENCE 2 (bases 1 to 1750)

AUTHORS Jakowlew,S.B.

TITLE Direct Submission

JOURNAL Submitted (14-JUL-1988) Jakowlew S.B., National Institute of
health, National Cancer Institute, Laboratory of Chemoprevention,
Building 41, Room B902, Bethesda, Maryland 20892, USA

COMMENT The submitters believe that the chicken cDNA library was
contaminated with porcine cDNA, and that the sequence is infact
porcine TGF-beta-1. 27-MAR-1996.

FEATURES Location/Qualifiers

source

1. .1750
/organism="Sus scrofa"
/mol_type="mRNA"
/strain="white leghorn"
/db_xref="taxon:9823"
/clone="pTGFB-ChX119"
/cell_type="chondrocyte"

gene

1. .1750
/gene="TGF-beta 1"

5'UTR

1. .446
/gene="TGF-beta 1"

CDS

447. .1622
/gene="TGF-beta 1"
/codon_start=1
/product="transforming growth factor"
/protein_id="CAA30933.1"

```

/db_xref="GI:63809"
/db_xref="SWISS-PROT:P07200"
/translation="MPPSGPGLLP LLLP LLLWLLVLT PGRPAAGLSTCKTIDMELVKRK
RIEAI RQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD
YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNDSWGYLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR
LSAHCSCDSKDNTLHVEINAGFNSGRRGDLATIHGMNRPFLLLMATPLERAQHLHSSR
HRRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD
TQYSKVLALYNQHNP GASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
mat_peptide 1285. .1619
3'UTR 1620. .1750
/gene="TGF-beta 1"
/gene="TGF-beta 1"
BASE COUNT 325 a 627 c 479 g 319 t
ORIGIN

```

```

Query Match 90.0%; Score 1213.8; DB 4; Length 1750;
Best Local Similarity 95.6%; Pred. No. 1.1e-204;
Matches 1286; Conservative 0; Mismatches 17; Indels 42; Gaps 2;

```

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     442 CCCCCATGCCGCCTTCGGGGCCTGGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 501

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     502 TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG 561

Qy     126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     562 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 621

Qy     186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 245
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     622 TCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 681

Qy     246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCTGAACCGGAGCCCGAGC 305
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     682 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCTGAACCGGAGCCCGAGC 741

Qy     306 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC 365
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     742 CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC 801

Qy     366 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     802 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 861

Qy     426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     862 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 921

Qy     486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT 545
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     922 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT 981

Qy     546 GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG 605

```

Db	982	 GGGGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCCTTTG	1041
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1042	 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	1101
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAAC--G	722
Db	1102	 TCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGCAG	1161
Qy	723	GTTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCC	782
Db	1162	 GTTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCC	1221
Qy	783	TGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCC	842
Db	1222	 TGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCC	1281
Qy	843	GAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACT	902
Db	1282	 GA-----GCCCTGGATACCAACTACT	1302
Qy	903	GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGG	962
Db	1303	 GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGG	1362
Qy	963	ACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGC	1022
Db	1363	 ACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGC	1422
Qy	1023	CCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACC	1082
Db	1423	 CCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACC	1482
Qy	1083	AGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGC	1142
Db	1483	 AGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGC	1542
Qy	1143	CCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1202
Db	1543	 CCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1602
Qy	1203	GTTCTTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCC	1262
Db	1603	 GTTCTTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCC	1662
Qy	1263	GGCCCCACCCCCGCCCGCTACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCAC	1322
Db	1663	 GGCCCCACCCCCGCCCGCTACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCAC	1722
Qy	1323	TTGGGATCGATTAAAGCGGCCGCGA	1347
Db	1723	 TTGGGATCGATTAAAGGTGGAGAGA	1747

Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTGTCCAAGCTTCGGC	185
Db	519	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTGTCCAAGCTTCGGC	578
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGG	245
Db	579	TCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGG	638
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAAGTGTGGAACCGGAGCCCGAGC	305
Db	639	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAAGTGTGGAACCGGAGCCCGAGC	698
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	699	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC	758
Qy	366	AAATCTATGATAAAATTCAGGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	759	AAATCTATGATAAAATTCAGGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	818
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	819	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	878
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	879	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	938
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	605
Db	939	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	998
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	999	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	1058
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1059	TCAGTGCCCACTGTTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	1118
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	785
Db	1119	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	1178
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1179	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGA-	1237
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1238	-----GCCCTGGATACCAACTACTGCT	1259
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1260	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	1319
Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025

```

      |||
Db      1320 TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT 1379
      |||
Qy      1026 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1085
      |||
Db      1380 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1439
      |||
Qy      1086 ACAACCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCCA 1145
      |||
Db      1440 ACAACCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCCA 1499
      |||
Qy      1146 TCGTGTA TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1205
      |||
Db      1500 TCGTGTA TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1559
      |||
Qy      1206 CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCAC 1251
      |||
Db      1560 CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCAC 1605

```

RESULT 5

AF461808

LOCUS AF461808 2221 bp mRNA linear MAM 03-JAN-2002

DEFINITION Sus scrofa transforming growth factor beta 1 (TGFB1) mRNA, complete cds.

ACCESSION AF461808

VERSION AF461808.1 GI:18042250

KEYWORDS .

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Polymorphism in the porcine transforming growth factor beta 1 gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2221)

AUTHORS Wimmers,K., Chomdej,S., Ponsuksili,S. and Schellander,K.

TITLE Direct Submission

JOURNAL Submitted (20-DEC-2001) Institute of Animal Breeding Science, University of Bonn, Endenicher Allee 15, Bonn 53115, Germany

FEATURES Location/Qualifiers

source 1. .2221
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/chromosome="6"

gene 1. .2221
/gene="TGFB1"

CDS 1. .1173
/gene="TGFB1"
/note="cytokine"
/codon_start=1
/product="transforming growth factor beta 1"
/protein_id="AAL57902.1"
/db_xref="GI:18042251"
/translation="MPPSGLRLLPLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK"

```

RIEAIRGQILSKLRLASPPSQGDVPPGPLPEAVLALYNSTRDRVAGESVEPEPEPEAD
YYAKEVTRVLMVESGNQIYDKFKGTPHSLYMLFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNDSWRYLSNRL LAPSDSPEWLSFDVTGVVRQWLTRREAIEGFR
LSAHCSCDSKDNTLHVEINGFNSGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNP GASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS"
1174. .2221
/gene="TGFB1"
3'UTR
BASE COUNT      504 a      616 c      649 g      452 t
ORIGIN

Query Match      83.2%; Score 1122.2; DB 4; Length 2221;
Best Local Similarity 96.2%; Pred. No. 1.6e-188;
Matches 1176; Conservative 0; Mismatches 8; Indels 39; Gaps 1;

Qy      11 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 70
      ||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGCCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 60

Qy      71 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 130
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 120

Qy      131 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 190
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTCGCC 180

Qy      191 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 250
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCCGTACTGGCTCTT 240

Qy      251 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 310
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 300

Qy      311 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACCAAATC 370
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACCAAATC 360

Qy      371 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 430
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 420

Qy      431 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 490
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 480

Qy      491 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC 550
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC 540

Qy      551 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC 610
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTC 600

Qy      611 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT 670
      |||||||||||||||||||||||||||||||||||||||||||||||||||

```



```

Db      601 ACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT 660
Qy      671 GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 730
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      661 GCCCACTGTTTCCTGTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 720
Qy      731 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 790
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      721 TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC 780
Qy      791 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG 850
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      781 ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGA----- 834
Qy      851 GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC 910
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      835 -----GCCCTGGATACCAACTACTGCTTCAGC 861
Qy      911 TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC 970
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      862 TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC 921
Qy      971 TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC 1030
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      922 TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC 981
Qy      1031 TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCTCTGTACAACCAGCACAAAC 1090
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      982 TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCTCTGTACAACCAGCACAAAC 1041
Qy      1091 CCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG 1150
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1042 CCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG 1101
Qy      1151 TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGC 1210
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1102 TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCCTTCCTGC 1161
Qy      1211 AAGTGCAGCTGAGGCCCCGCCCC 1233
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1162 AAGTGCAGCTGAGGCCCCAAGCC 1184

```

RESULT 6

DOGTGFB1A

LOCUS DOGTGFB1A 1369 bp mRNA linear MAM 30-OCT-1994

DEFINITION Canine transforming growth factor-beta 1 (TGFB1) mRNA, complete cds.

ACCESSION L34956

VERSION L34956.1 GI:516071

KEYWORDS homologue; transforming growth factor-beta 1.

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 1369)

AUTHORS Manning,A.M., Auchampach,J.A., Drong,R.F. and Slightom,J.L.

Db	233	TCTCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGTGCCGCTGCCCGAGGCCGTGCTGG	292
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCTGAACCGGAGCCCGAGC	305
Db	293	CCCTCTACAACAGCACCCGCGACCGGGTGGCGGGGGAGAGCGCCGAGCCGGAGCCCGAGC	352
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	353	CCGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAACACCAACA	412
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	413	AAATCTATGAGAAAGTCAAGAAAAGTCCGCACAGCATATATATGCTCTTCAACACATCAG	472
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	473	AGCTCCGAGAAGCAGTGCTGAGCCCGTCTTGCTCTCCCGGCAGAGTTGCGCCTGCTGA	532
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT	545
Db	533	GGCTCAAGTTAAAAGCGGAGCAGCATGTGGAGCTGTACCAGAAATATAGCAATGATTCTCT	592
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	593	GGCGCTACCTCAGCAACCGGCTGCTGGCGCCCAGCGACACGCCAGAATGGCTGTCCTTTG	652
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	653	ATGTCACTGGAGTCGTGAGGCAGTGGCTGAGCCATGGAGGGGAAGTCGAGGGCTTTTCGCC	712
Qy	666	TCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAAACGGGT	725
Db	713	TCAGTGCCCACTGTTCTCTGTGACAGCAAAGATAACACACTGCAAGTAGACATTAAACGGGT	772
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCTCTGC	785
Db	773	TCAGTTCAGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGACCCCTTCTCTGC	832
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	833	TCCTCATGGCCACCCCACTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCAGCGCCG--	890
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	891	-----GGCCCTGGACACCAACTACTGCT	913
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	914	TCAGCTCCACGGAGAAGAACTGCTGCGTCCGGCAGCTCTACATTGACTTCCGCAAGGATC	973
Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	974	TGGGCTGGAAGTGGATCCATGAGCCCAAGGGTTACCACGCTAATTCTGCCTGGGGCCCT	1033
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085

```

Db      1034  GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 1093
Qy      1086  ACAACCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCCA 1145
        |||
Db      1094  ACAACCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCCA 1153
        |||
Qy      1146  TCGTGTAACGCTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1205
        |||
Db      1154  TCGTGTAACGCTGGGCGCAAGCCCAAGGTGGAGCAGCTGTGGAACATGATCGTGCGCT 1213
        |||
Qy      1206  CCTGCAAGTGCAGCTGAGGCCCCGCCCCG-CCCACAGCCCCGCCACCCGGCAGGCCCCGG 1264
        |||
Db      1214  CCTGCAAGTGCAGCTGAGGCCCCGCCCCGTCCGGCAGGCCCCGCCACCCGGCAGGNCCGG 1273
        |||
Qy      1265  CCCCACCCCGCCCGCTCACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGC 1318
        |||
Db      1274  CCCCACCCCGCCCGCTGCGCC-GGGCTGTATTTAAGGACACCGCGCCCCAGGC 1326
        |||

```

RESULT 7

E00973

LOCUS E00973 2527 bp RNA linear PAT 29-SEP-1997

DEFINITION cDNA encoding human TGF-beta.

ACCESSION E00973

VERSION E00973.1 GI:2169234

KEYWORDS JP 1986219395-A/1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2527)

AUTHORS Riku,M.A.D.D. and Debitsudo,B.G.

TITLE NUCLEIC ACID CODED WITH TGF-BETA AND ITS USE

JOURNAL Patent: JP 1986219395-A 1 29-SEP-1986;

GENENTECH INC

COMMENT

OS human

PN JP 1986219395-A/1

PD 29-SEP-1986

PF 20-MAR-1986 JP 1986064661

PR 22-MAR-1985 US 85 715142

PI RIKU MAIKERU ANDORE DERINKU, DEBITSUDO BANNOOMAN GETSUDERU PC
C12P21/00,C12N1/00,C12N5/00,C12N15/00//C12Q1/68,(C12P21/00, PC
C12R1:91),

PC (C12N1/00,C12R1:19),(C12N5/00,C12R1:91),(C12N15/00,C12R1:91);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC *source: tissue_type=placenta and glyoblastoma; FH Key
Location/Qualifiers

FH

FT CDS 842. .2014

FT /product='pre TGF-beta'

FT mat_peptide 1676. .2011

FT /product='TGF-beta'

FT 5'UTR 1. .841

FT 3'UTR 2015. .2537

```

FT      stem_loop      37. .113
FT      stem_loop      2015. .2100
FT      polyA_site      2514. .2519.
FEATURES                      Location/Qualifiers
  source                      1. .2527
                              /organism="Homo sapiens"
                              /mol_type="genomic RNA"
                              /db_xref="taxon:9606"
BASE COUNT      472 a      888 c      735 g      432 t
ORIGIN

```

```

Query Match      73.5%; Score 991.6; DB 6; Length 2527;
Best Local Similarity 85.7%; Pred. No. 1.9e-165;
Matches 1142; Conservative 0; Mismatches 149; Indels 41; Gaps 2;

```

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGGCTGCTGTGGCTGC 65
      || ||| ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      897 TGGTGCTGACGCCTGGCCCGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy      186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 245
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCCAGGCCGTGCTCG 1076

Qy      246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGGAGCCCGAGC 305
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1077 CCCTCTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAAACCGGAGCCCGAGC 1136

Qy      306 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 365
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1137 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG 1196

Qy      366 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy      426 AGCTCCGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1257 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 1316

Qy      486 GGCTCAAGTTAAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1317 GGCTCAAGTTAAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT 1376

Qy      546 GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTTG 605
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1377 GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGTTATCTTTTG 1436

Qy      606 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGAGGCTATAGAGGGTTTTCGCC 665
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


KEYWORDS .
 SOURCE synthetic construct
 ORGANISM synthetic construct
 artificial sequences.
 REFERENCE 1 (bases 1 to 2537)
 AUTHORS .
 TITLE NUCLEIC ACID ENCODING TGF- beta 3 AND ITS USE
 JOURNAL Patent: WO 8912101-A 4 14-DEC-1989;
 FEATURES Location/Qualifiers
 source 1. .2537
 /organism="synthetic construct"
 /mol_type="mRNA"
 /db_xref="taxon:32630"
 CDS 842. .2014
 /codon_start=1
 /transl_table=11
 /product="preTGF-beta1"
 /protein_id="CAA00588.1"
 /db_xref="GI:412941"
 /translation="MPPSGLRLLPLLLPLLWLLVLTGPPPAAGLSTCKTIDMELVKRK
 RIEAIRGQILSKRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
 YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRL
 KLKVEQHVELYQKYSNNSWRYLSNRL LAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFR
 LSAHCSCDSRDNTLQVDINGFTTGRRGDLATIHGMRPFLLLMATPLERAQHLQSSRH
 RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
 QYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
 BASE COUNT 473 a 893 c 739 g 432 t
 ORIGIN

Query Match 73.4%; Score 990.8; DB 6; Length 2537;
 Best Local Similarity 85.5%; Pred. No. 2.6e-165;
 Matches 1146; Conservative 0; Mismatches 147; Indels 47; Gaps 2;

Qy	6	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCGCTGCTGTGGCTGC	65
Db	837	CCCCCATGCGGCCCTCCGGGCTGCGGCTGCTGCGCTGCTGCTACCGCTGCTGTGGCTAC	896
Qy	66	TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	897	TGGTGCTGACGCCTGGCCCGCCGGCCGGCGGACTATCCACCTGCAAGACTATCGACATGG	956
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	957	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	1016
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG	1196

Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1376
Qy	546	GGCGCTACCTCAGCAACCGGTGCTGGCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGTGCTGGCACCCAGCGACTCGCCAGAGTGTTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGCGTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGCGTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCTCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1758	TGGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT	1817
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1818	GCCCCATCATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877
Qy	1086	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1145
Db	1878	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA	1937
Qy	1146	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1205
Db	1938	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1997
Qy	1206	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC	1265


```

Db      1998 CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGGCAGGCCCGGCCCCAC 2057
Qy      1266 CCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG 1317
Db      2058 CCCGCCCCGCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG 2117
Qy      1318 CCCACTTGGGATCGATTAAA 1337
Db      2118 CCCACCTGGGGCCCCATTAA 2137

```

RESULT 9

OATGFB1

LOCUS	OATGFB1	1173 bp	mRNA	linear	MAM 18-APR-1995
-------	---------	---------	------	--------	-----------------

DEFINITION O.aries mRNA for transforming growth factor-beta 1.

ACCESSION X76916

VERSION X76916.1 GI:496648

KEYWORDS TGF-beta 1; transforming growth factor-beta 1.

SOURCE Ovis aries (sheep)

ORGANISM *Ovis aries*

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Caprinae; Ovis.

REFERENCE 1

AUTHORS Woodall, C.J., McLaren, L.J. and Watt, N.J.

TITLE Sequence and chromosomal localisation of the gene encoding ovine
 latent transforming growth factor-beta 1

JOURNAL Gene 150 (2), 371-373 (1994)

MEDLINE 95121932

PUBMED 7821809

REFERENCE 2 (bases 1 to 1173)

AUTHORS Woodall, C.

TITLE Direct Submission

JOURNAL Submitted (24-DEC-1993) C. Woodall, Univ. of Edinburgh, Dept. of
Veterinary Pathology, Sc. of Vet. Studies, Univ. of Edinburgh,
Edinburgh EH9 1QH, UK

FEATURES	Location/Qualifiers
----------	---------------------

```
source      1.  .1173
```

```
/organism="Ovis aries"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9940"
```

CDS 1. $\bar{.1173}$

```
/codon start=1
```

```
/product="transforming growth factor-beta 1"
```

```
/protein id="CAA54242.1"
```

```
/db_xref="GI:496649"
```

```
/db_xref="SWISS-PROT:P50414"
```

```
/translation="MPPSGLRLLPLLLPLLWLLMLTPGRP VAGLSTCKTIDMELVKRK
```

GIEAIRGQILSKLRLASPPSQGDVPPGPLPEAILALYNSTRDRVAGESAETEPEPEAD

YYAKEVTRVLMVEYGNKIYDKMKSSSHSIYMFFNTSELREAVPEPVLLSRADVRLRL

KLKVEQHVELYQKYSNNSWRYLSNRL LAPSDSPEWLSFDVTGVVRQWLTHREEIEGFR

LSAHCSCDSKDNTLQVDINGFSSGRRGDLATI HGMNRPFLLLMATPLERAQHLHSSRH

RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLD

mat peptide	835.1170
-------------	----------

```
/product="transforming growth factor-beta 1"
```

BASE COUNT	245 a	378 c	336 g	214 t
ORIGIN				

Query Match 73.0%; Score 984.8; DB 4; Length 1173;
Best Local Similarity 89.6%; Pred. No. 3.3e-164;
Matches 1086; Conservative 0; Mismatches 87; Indels 39; Gaps 1;

[illegible]

Db	721	TCCGGCCGCGGGGTGACCTCGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	780
Qy	791	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCCGCCACCGCCGAGCCCTG	850
Db	781	ATGGCCACCCCTCTGGAGAGGGCCCAGCACCTGCACAGCTCCCCGCCACCGCCGA-----	834
Qy	851	GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC	910
Db	835	-----GCCCTGGACACCAACTACTGCTTCAGC	861
Qy	911	TCCACGGAGAAGAACTGCTGCGTGC GG CAGCTCTACATTGACTTCCGGAAGGACCTGGGC	970
Db	862	TCCACAGAAAAGAACTGCTGTGTTCTGTCAGCTCTACATTGACTTCCGGAAGGACCTGGGC	921
Qy	971	TGGAAGTGGAATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1030
Db	922	TGGAAGTGGAATTCACGAACCCAAGGGCTACCACGCCAATTTCTGCCTGGGGCCCTGTCCC	981
Qy	1031	TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAAAC	1090
Db	982	TACATCTGGAGCCTGGACACACAGTACAGCAAGGTCTGGCCCTGTACAACCAGCACAAAC	1041
Qy	1091	CCGGGCGCGTCTGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1150
Db	1042	CCGGGCGCATCGGCGGGCGCCGTGCTGCGTGCCCTCAGGCGCTGGAACCCCTGCCCATCGTG	1101
Qy	1151	TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC	1210
Db	1102	TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGTTGTCCAACATGATCGTGCGTTCCTGC	1161
Qy	1211	AAGTGCAGCTGA 1222	
Db	1162	AAGTGCAGCTGA 1173	

Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 6 Row: e Column: 11
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES	Location/Qualifiers
source	1. .1780 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:3119 IMAGE:3351664" /tissue_type="Eye, retinoblastoma" /clone_lib="NIH_MGC_16" /lab_host="DH10B-R" /note="Vector: pOTB7"
CDS	447. .1619 /codon_start=1 /product="Similar to transforming growth factor, beta 1" /protein_id="AAH00125.1" /db_xref="GI:12652749" /translation="MPPSGLRLLLLLLLPLLWLLVLTTPGRPAAGLSTCKTIDMELVKRK RIEAIRGQILSKLRLASFPSPQGEVPPGPPLPEAVLALYNSTRDRVAGESAEPEPEPEAD YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFNTSELREAVPEPVLLSRAELRLLRL KLKVEQHVELYQKYSNNSWRYLNSRLLAPSDSPEWLSFDVTGVVRQWLRSRGGEIEGFR LSAHCSDSRDNTLQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT QYSKVLALYNQHNPASAAPCCVPQALEPLIVYYVGRKPKVEQLSNMIVRSCKCS"
BASE COUNT	341 a 648 c 493 g 298 t
ORIGIN	

Query Match 73.0%; Score 984.4; DB 9; Length 1780;
 Best Local Similarity 85.2%; Pred. No. 3.7e-164;
 Matches 1142; Conservative 0; Mismatches 151; Indels 47; Gaps 2;

Qy	6	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC	65
Db	442	CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTAC	501
Qy	66	TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	502	TGGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTATCCACCTGCAAGACTATCGACATGG	561

Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	562	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATCCTGTCCAAGCTGCGGC	621
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	622	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG	681
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGC	305
Db	682	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	741
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGAAAGCGGCAACC	365
Db	742	CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGAAACCCACAACG	801
Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	802	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	861
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	862	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	921
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT	545
Db	922	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT	981
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTTG	605
Db	982	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1041
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1042	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1101
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1102	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGACATCAACGGGT	1161
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGGCATGAACCGGCCCTTCTGC	785
Db	1162	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGC	1221
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1222	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1280
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1281	-----GCCCTGGACACCAACTATTGCT	1302
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1303	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1362

Db	622 TCGCCAGCCCCCAGGCCAGGGGAGGTGCCGCCCGGCCCTGCCCGAGGCCGTGCTCG	681
Qy	246 CTCTTTACAACAGTACCCGCGACC GG GTAGCCGGGGAAAGTGTCGAACCGGAGCCC GAGC	305
Db	682 CCCTGTACAACAGCACCCGCGACC GG GTGGCCGGGGAGAGTG CAGA ACCGGAGCCC GAGC	741
Qy	306 CAGAGGCGGACTACTACGCCAAGGAGGTCA CCC CGCTGC TAATGGTGGA AAGCGG CAACC	365
Db	742 CTGAGGCCGACTACTACGCCAAGGAGGTCA CCC CGCTGC TAATGGTGGA AACCCA ACAG	801
Qy	366 AAATCTATGATAAA TTCAAGGGCACCCCCAC AGCTTATATATGCTGTTCAACACGTCGG	425
Db	802 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	861
Qy	426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	862 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	921
Qy	486 GGCTCAAGTTAAAAGTG GAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATT CCT	545
Db	922 GGCTCAAGTTAAAAGTG GAGCAGCACGTGGAGCTGTACCAGAAATACAGCAA CAAT TCCT	981
Qy	546 GCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTG ACTCACCGGAGTG GCTGTCCTTTG	605
Db	982 GCGGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTG GTTATCTTTTG	1041
Qy	606 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCG CAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1042 ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTG GAGGGGAAATTGAGGGCTTTTCGCC	1101
Qy	666 TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1102 TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGA CATCAACGGGT	1161
Qy	726 TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTCA CGGCATGAACCGGCCCTTCCTGC	785
Db	1162 TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTCA TG GCATGAACCGGCCCTTCCTGC	1221
Qy	786 TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1222 TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1280
Qy	846 CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1281 -----GCCCTGGACACCAACTATTGCT	1302
Qy	906 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1303 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1362
Qy	966 TGGGCTGGAAGTGGA TTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1363 TCGGCTGGAAGTGGA TCACAGAGCCCAAGGGCTACCATGCCA ACTTCTGCCTCGGGCCCT	1422
Qy	1026 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1423 GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1482

[illegible]

RESULT 12

```

AGMTGFB
LOCUS      AGMTGFB                      1561 bp    mRNA    linear    PRI 27-APR-1993
DEFINITION Simian transforming growth factor-beta (TGF) mRNA, complete cds.
ACCESSION  M16658
VERSION    M16658.1  GI:176552
KEYWORDS   growth factor; transforming growth factor-beta.
SOURCE     Cercopithecus aethiops (African green monkey)
  ORGANISM Cercopithecus aethiops
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Cercopithecus.
REFERENCE  1 (bases 1 to 1561)
  AUTHORS  Sharples,K., Plowman,G.D., Rose,T.M., Twardzik,D.R. and
            Purchio,A.F.
  TITLE    Cloning and sequence analysis of simian transforming growth
            factor-beta cDNA
  JOURNAL  DNA 6 (3), 239-244 (1987)
  MEDLINE  87246074
  PUBMED   3474130
COMMENT    Original source text: African green monkey cells (cell line
            BSC-40), cDNA to mRNA, clone pTGF-beta-2.
FEATURES   Location/Qualifiers
    source          1. .1561
                   /organism="Cercopithecus aethiops"
                   /mol_type="mRNA"
                   /db_xref="taxon:9534"
    CDS            262. .1434
                   /note="transforming growth factor-beta precursor"
                   /codon_start=1
                   /protein_id="AAA35369.1"
                   /db_xref="GI:176553"
                   /translation="MPPSGLRLLPLLLPLLWLVLVTPSRPAAGLSTCKTIDMELVKRK
            RIETIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
            YYAKEVTRVLMVETHNEIYDKFKOSTHSIYMFFNTSELREAVPEPVLLSRAELRLRL

```

```

KLKVEQHVELYQKYSNNSWRYLSNRLLAPSNPEWLSFDVTGVVRQWLSRGGEIEGFR
LSAHCSCDSKDNLTQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPASAAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
sig_peptide      262. .1095
                  /note="transforming growth factor-beta signal peptide"
mat_peptide      1096. .1431
                  /product="transforming growth factor-beta"
BASE COUNT      301 a      547 c      446 g      267 t
ORIGIN

```

```

Query Match          72.9%; Score 983; DB 9; Length 1561;
Best Local Similarity 85.6%; Pred. No. 6.6e-164;
Matches 1148; Conservative 0; Mismatches 145; Indels 48; Gaps 3;

```

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      || ||| ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     257 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 316

Qy     66 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    317 TGGTGCTGACGCCTAGCCGGCCGGCCGGCCGAGACTATCCACCTGCAAGACTATCGACATGG 376

Qy    126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    377 AGCTGGTGAAGCGGAAGCGCATCGAGACCATCCGGGCCAGATCCTGTCCAAGCTGCGGC 436

Qy    186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    437 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 496

Qy    246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    497 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGGAGCCGGAGCCCGAAC 556

Qy    306 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 365
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    557 CGGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG 616

Qy    366 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    617 AAATCTATGACAAGTTCAAGCAGAGCACACAGCATATATATGTTCTTCAACACATCAG 676

Qy    426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    677 AGCTCCGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCGGGCAGAGCTGCGTCTGCTGA 736

Qy    486 GGCTCAAGTTAAAAGTGGAGCAGCAGTGGAGCTATACCAGAAATACAGCAATGATTCTCT 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    737 GGCTCAAGTTAAAAGTGGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCCT 796

Qy    546 GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCTTTTG 605
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    797 GGCGATACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGAGTGGTTGTCTTTTG 856

Qy    606 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC 665
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	857	ATGTCACCCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCC	916
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAAACGGGT	725
Db	917	TTAGCGCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGT	976
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	977	TCACTACCGGCCGCCGAGGTGACCTGGCCACAATTATGTCATGAACCGGCCCTTCTCTGC	1036
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1037	TTCTCATGGCCACCCCACTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-	1095
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1096	-----GCCCTGGACACCAACTACTGCT	1117
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1118	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACC	1177
Qy	966	TGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1178	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTGGGGCCCT	1237
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1238	GTCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1297
Qy	1086	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1145
Db	1298	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1357
Qy	1146	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1205
Db	1358	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1417
Qy	1206	CCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC	1265
Db	1418	CCTGCAAATGCAGCTGAGGCCCGCCCCGCCCGCCCCACCCCGGCAGGCCCGGCCCGC	1477
Qy	1266	CCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAA	1316
Db	1478	CCCACCCCAACCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTGCCCCAA	1537
Qy	1317	GCCCACTTGGGATCGATTAAA	1337
Db	1538	GCCCACTTGGGGCCCCATTAA	1558

Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	627	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	686
Qy	186	TTGCCAGCCCCCGAGCCAGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	687	TCGCCAGCCCCCGAGCCAGGGGAGGTGCCGCCCGGCCCGCTGCCCCGAGGCCGTGCTCG	746
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGAAAAGTGTGCAACCGGAGCCCGAGC	305
Db	747	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	806
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	807	CTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAACCCACAACG	866
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTGG	425
Db	867	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	926
Qy	426	AGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	986
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	987	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1046
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	605
Db	1047	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGCGTTATCTTTG	1106
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1166
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1167	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGACATCAACGGGT	1226
Qy	726	TCAATTCTGGCCGCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGTC	785
Db	1227	TCACTACCGGCCGCGGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTGTC	1286
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1346	-----GCCCTGGACACCAACTATTGCT	1367
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1368	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1427

Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1428	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCT	1487
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1488	GCCCTTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1547
Qy	1086	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA	1145
Db	1548	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCA	1607
Qy	1146	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1205
Db	1608	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1667
Qy	1206	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC	1265
Db	1668	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCCAGGCCCCGGCCCCAC	1727
Qy	1266	CCCACCCCCGCCCCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1317
Db	1728	CCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAA	1787
Qy	1318	CCC 1320	
Db	1788	GCC 1790	

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 27 Row: e Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

```
FEATURES             Location/Qualifiers
     source            1..1746
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="MGC:22008 IMAGE:4399762"
                        /tissue_type="Duodenum, adenocarcinoma"
                        /clone_lib="NIH_MGC_88"
                        /lab_host="DH10B"
                        /note="Vector: pCMV-SPORT6"
     CDS               370..1542
                        /codon_start=1
                        /product="Unknown (protein for MGC:22008)"
                        /protein_id="AAH22242.1"
                        /db_xref="GI:18490116"
                        /translation="MPPSGLRLLLLLLPLLWLLVLTGPRPAAGLSTCKTIDMELVKRK
RIEAIIRGQILSKLRLASPPSQGEVPPGPLPEAVLALYNSTRDRVAGESAEPEPEPEAD
YYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFFNTSELREAVPEPVLLSRAELRLLRL
KLKVEQHVELYQKYSNNSWRYLSNRLAPSDSPEWLSFDVTGVVRQWLSRGGEIEGFR
LSAHCSCDSRDNTLQVDINGFTTGRRGDLATIHGMNRPFLLLMATPLERAQHLQSSRH
RRALDTNYCFSSTEKNCCVRQLYIDFRKDLGWKWIHEPKGYHANFCLGPCPYIWSLDT
QYSKVLALYNQHNPGASAAPCCVPQALEPLPIVYYVGRKPKVEQLSNMIVRSCKCS"
```

BASE COUNT 376 a 612 c 472 g 286 t
ORIGIN

Query Match 72.6%; Score 979.8; DB 9; Length 1746;
Best Local Similarity 85.5%; Pred. No. 2.4e-163;
Matches 1146; Conservative 0; Mismatches 147; Indels 48; Gaps 3;

```
Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     365 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTAC 424

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      |  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     425 TGGTGCTGACGCCTGGCCGGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG 484

Qy     126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     485 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCCGCGGCCAGATCCTGTCCAAGCTGCGGC 544

Qy     186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      |  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     545 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 604

Qy     246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      ||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     605 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 664
```

Qy	306	CAGAGCGCGGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	665	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAACCCACAACG	724
Qy	366	AAATCTATGATAAAATTCAGGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	725	AAATCTATGACAAGTTCAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	784
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	785	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	844
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCCT	545
Db	845	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCCT	904
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	905	GGCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTG	964
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	965	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1024
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1025	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1084
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCCTGC	785
Db	1085	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTTCTGC	1144
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1145	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1203
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1204	-----GCCCTGGACACCAACTATTGCT	1225
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1226	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1285
Qy	966	TGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1286	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT	1345
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1346	GCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1405
Qy	1086	ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1145
Db	1406	ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA	1465

Qy	1146	TCGTGTA	CTACGTGGGCCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1205
Db	1466	TCGTGTA	CTACGTGGGCCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1525
Qy	1206	CCTGCAAGTGCAGCTGAGGCCCCGCCCGCCCACAGCCCCGCCCACCCGGCAGGCCCCGGC	1265	
Db	1526	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCAGGCCCCGGCCCCAC	1585	
Qy	1266	CCCACCCCCGCCCCGCT-----CACCGGGGCTGTATT	TAAAGGACA-TCGTGCCCCAA	1316
Db	1586	CCCGCCCCGCCCCGCTGCGCTTGCCCATGGGGGCTGTATT	TAAAGGACACCCGTGCCCCAA	1645
Qy	1317	GCCCACTTGGGATCGATTAAA	1337	
Db	1646	GCCCACTTGGGGCCCCATTAA	1666	

Query Match 72.5%; Score 978.6; DB 6; Length 1560;
Best Local Similarity 85.6%; Pred. No. 4e-163;
Matches 1143; Conservative 0; Mismatches 145; Indels 48; Gaps 3;

Qy	251	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAAGTGTGCAACCCGGAGCCCGAGCCAGAG	310
Db	501	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCGGAGCCCGAGCCCGAACCGGAG	560
Qy	311	GCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	370
Db	561	GCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACGAAATC	620
Qy	371	TATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTTCGGAGCTC	430
Db	621	TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	431	CGGGAAGCGGTGCGCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	490
Db	681	CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	491	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	550
Db	741	AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	800
Qy	551	TACCTCAGCAACCGGTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTC	610
Db	801	TACCTCAGCAACCGGTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	611	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	670
Db	861	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	920
Qy	671	GCCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	730
Db	921	GCCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	980
Qy	731	TCTGGCCGCGCGGGGTGACCTGGCCACCATTACAGGCATGAACCGGCCCTTCCTGCTCCTC	790
Db	981	ACCGGCCGCGGAGGTGACCTGGCCACAATTGATGGCATGAACCGGCCCTTCCTGCTTCTC	1040
Qy	791	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	850
Db	1041	ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-----	1094
Qy	851	GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC	910
Db	1095	-----GCCCTGGACACCAACTACTGCTTCAGC	1121
Qy	911	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC	970
Db	1122	TCCACGGAGAAGAACTGCTNCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGC	1181
Qy	971	TGGAAGTGGATTTCATGAACCCAAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1030
Db	1182	TGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1241
Qy	1031	TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAAAC	1090
Db	1242	TACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAAC	1301
Qy	1091	CCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1150

Db	1302	CCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1361
Qy	1151	TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC	1210
Db	1362	TACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGA	1421
Qy	1211	AAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGCCCCAC	1270
Db	1422	AAATGCAGCTGAGGCCCCGCCCCGCCCGCCCCACCCCGGCAGGCCCCGGCCCCGCCCCAC	1481
Qy	1271	CCCCGCCGCCT-----CACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAAGCCCA	1321
Db	1482	CCCACCCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTGCCCCAAGCCCA	1541
Qy	1322	CTTGGGATCGATTAAA	1337
Db	1542	CCTGGGGCCCCATTAA	1557

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

```
Run on:      October 27, 2003, 10:47:27 ; Search time 405.282 Seconds
              (without alignments)
              8985.201 Million cell updates/sec
```

6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	1223.8	90.7	1326	24	AAD22696	Porcine transformi
2	991.6	73.5	2527	25	ABQ76674	Androgen receptor
3	990.8	73.4	2537	7	AAN60972	Sequence encoding
4	990.8	73.4	2537	11	AAQ03301	cDNA encoding huma
5	990.8	73.4	2537	11	AAQ02814	Sequence of pre-TG
6	990.8	73.4	2537	17	AAT15720	Pre-transforming g
7	989.2	73.3	2537	15	AAQ56923	Human pre-TGF-beta
8	986	73.1	2537	19	AAV52933	Human pre-transfor
9	983	72.9	1561	11	AAQ03268	Simian transformin
10	981.4	72.8	2742	22	AAI58342	Human polynucleoti
11	980.6	72.7	1559	13	AAQ20289	Sequence encoding
12	980.2	72.7	1821	12	AAQ13392	Human pro-TGF-beta
13	979.6	72.6	1560	9	AAN81084	Coding sequence of
14	979.6	72.6	1560	11	AAQ03508	Simian Transformin
15	977.8	72.5	2745	16	AAT05876	cDNA encoding tran
16	977.8	72.5	2745	22	AAH28216	Nucleotide sequenc
17	969.8	71.9	1303	11	AAQ09317	Monkey transformin
18	968.8	71.8	4105	15	AAQ55624	TGFbeta1 5'-UTR-CD
19	966.8	71.7	1571	11	AAQ03269	Human transforming
20	961.8	71.3	1569	9	AAN81085	Coding sequence of
21	958.6	71.1	1569	11	AAQ03509	Human Transforming
22	952.6	70.6	1561	11	AAQ04908	Sequence encoding
23	949.4	70.4	1561	13	AAQ29177	TGF-beta 1/beta 2
24	930.2	69.0	1176	25	ABV75391	TGFB1 Arg25Pro pol
25	928.6	68.8	1176	24	ABZ35738	Human TGF beta 1 p
26	928.6	68.8	1176	24	ABX09981	Human TGFbeta1 DNA
27	928.6	68.8	1176	24	ABV78162	Human TGF beta 1 D

	28	928.6	68.8	1176	24	ABL91703	Human polynucleoti
	29	928.6	68.8	1176	25	ABV75392	TGFB1 Arg25Pro pol
	30	908	67.3	1565	13	AAQ29178	TGF-beta 1. Homo
c	31	805	59.7	2765	22	AAI60128	Human polynucleoti
	32	788.2	58.4	2208	13	AAQ20291	Sequence encoding
	33	786.6	58.3	2206	11	AAQ03510	Human Transforming
	34	786.6	58.3	2207	11	AAQ03511	Hybrid transformin
	35	785	58.2	2207	11	AAQ05127	Human TGF-Betal/TG
	36	773.6	57.3	2217	10	AAN90768	Sequence of human
	37	767.4	56.9	2200	16	AAT04115	Simian-human hybri
c	38	696.8	51.7	2773	23	AAS84421	DNA encoding novel
	39	678.8	50.3	834	12	AAQ12192	Sequence encoding
	40	654	48.5	1376	24	ABK90341	DNA encoding LAP-m
	41	652.4	48.4	1389	24	ABK90344	DNA encoding LAP-h
	42	588.2	43.6	1352	24	ABK90342	DNA encoding mIFNB
	43	586.6	43.5	1350	24	ABK90343	DNA encoding huIFN
	44	534.8	39.6	875	23	AAS70979	DNA encoding novel
	45	366.4	27.2	489	24	ABL99528	Target canine gene

ALIGNMENTS

RESULT 1

AAD22696

ID AAD22696 standard; cDNA; 1326 BP.

XX

AC AAD22696;

XX

DT 26-FEB-2002 (first entry)

XX

DE Porcine transforming growth factor beta 1 (TGF-betal) cDNA.

XX

KW Porcine; transforming growth factor beta 1; TGF-betal; gene therapy;
 KW IBD; inflammatory bowel disease; autoimmune disease; immunosuppressive;
 KW multiple sclerosis; rheumatoid arthritis; systemic lupus erythematosus;
 KW diabetes mellitus; sarcoidosis; psoriasis; dermatological; ss.

XX

OS Sus scrofa.

XX

FH Key Location/Qualifiers

FT CDS 16..1188

FT /*tag= a

FT /product= "Porcine TGF-betal mutant protein"

XX

PN WO200181404-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12980.

XX

PR 20-APR-2000; 2000US-199014P.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Strober W, Nakamura K, Kitani A, Fuss IJ;

XX

DR WPI; 2002-026155/03.

DR P-PSDB; AAE13596.

XX

PT Composition for treating autoimmune diseases e.g. inflammatory bowel
PT disease in humans, comprises vector containing transforming growth
PT factor-beta under the control of inducible promoter -

XX

PS Claim 1; Fig 1; 78pp; English.

XX

CC The invention relates to a composition containing a vector comprising a
CC gene encoding a regulatory transcription factor under the control of a
CC promoter encoding a transforming growth factor-beta (TGF-beta). The
CC vector is useful for expressing TGF-beta, such as TGF-beta1, TGF-beta2
CC or TGF-beta3, its variants or homologues, by transfecting a cell which
CC is part of a host suspected of having an autoimmune disease, especially
CC inflammatory bowel disease (IBD), under conditions such that the
CC polypeptide encoded by the nucleic acid sequence in the vector is
CC expressed. The vector is delivered using a delivery system. The delivery
CC of the vector results in substantial elimination of symptoms of the
CC autoimmune disease and increased production of IL-10 by the host. The
CC composition is useful for treating various diseases with an autoimmune
CC component such as multiple sclerosis, rheumatoid arthritis, systemic
CC lupus erythematosus, insulin-dependent diabetes mellitus, sarcoidosis
CC and psoriasis, and also for assaying the expression of a gene in a cell.
CC The vector is further useful for screening of the effect of test
CC compounds on cytokine (e.g. TGF-beta) expression of transfected cells.
CC The present sequence is a cDNA encoding porcine TGF-beta1 mutant.

XX

SQ Sequence 1326 BP; 263 A; 438 C; 392 G; 233 T; 0 other;

Query Match 90.7%; Score 1223.8; DB 24; Length 1326;
Best Local Similarity 96.4%; Pred. No. 1.1e-240;
Matches 1299; Conservative 0; Mismatches 7; Indels 41; Gaps 3;

Qy	1	TGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTG	60
Db	6	TGGTACCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTG	65
Qy	61	GCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGA	120
Db	66	GCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGA	125
Qy	121	CATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCT	180
Db	126	CATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCT	185
Qy	181	TCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGT	240
Db	186	TCGGCTCGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCCGT	245
Qy	241	ACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCC	300
Db	246	ACTGGCTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCC	305
Qy	301	CGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGG	360
Db	306	CGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGG	365

Qy	361	CAACCAAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACAC	420
Db	366	CAACCAAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACAC	425
Qy	421	GTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCT	480
Db	426	GTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCT	485
Qy	481	GCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGA	540
Db	486	GCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGA	545
Qy	541	TTCCTGGCGCTACCTCAGCAACCGGTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTC	600
Db	546	TTCCTGGCGCTACCTCAGCAACCGGTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTC	605
Qy	601	CTTTGATGTCAACCGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTT	660
Db	606	CTTTGATGTCAACCGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTT	665
Qy	661	TCGCCTCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCAGTGGAAATTAA	720
Db	666	TCGCCTCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCAGTGGAAATTAA	725
Qy	721	CGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTT	780
Db	726	CGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTT	785
Qy	781	CCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCG	840
Db	786	CCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCG	845
Qy	841	CCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTA	900
Db	846	CCGA-----GCCCTGGATACCAACTA	866
Qy	901	CTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAA	960
Db	867	CTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAA	926
Qy	961	GGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGG	1020
Db	927	GGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGG	986
Qy	1021	GCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAA	1080
Db	987	GCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAA	1046
Qy	1081	CCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACT	1140
Db	1047	CCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACT	1106
Qy	1141	GCCCATCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGT	1200
Db	1107	GCCCATCGTGTA CTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGT	1166


```
AC      AAN60972;
XX
DT      31-OCT-2002   (updated)
DT      28-OCT-1991   (first entry)
XX
DE      Sequence encoding preTGF-beta.
XX
KW      Transforming growth factor beta; cancer; wound healing.
XX
OS      Unidentified.
XX
FH      Key           Location/Qualifiers
FT      misc_structure 37..113
FT                               /*tag=  a
FT                               /note= "Sequence can form stable hairpin loops"
FT      CDS            842..2014
FT                               /*tag=  b
FT      mat peptide    1676..2011
FT                               /*tag=  c
XX
PN      EP200341-A.
XX
PD      10-DEC-1986.
XX
PF      21-MAR-1986;   86EP-0302112.
XX
PR      22-MAR-1985;   85US-0715142.
PR      13-MAR-1987;   87US-0025423.
XX
PA      (GETH ) GENENTECH INC.
XX
PI      Derynck RMA;
XX
DR      WPI; 1986-326875/50.
DR      P-PSDB; AAP61468.
XX
PT      TGF-beta prodn. from transformed hosts - useful esp. for treating
PT      wounds (J6 2/9/86).
XX
PS      Disclosure; Fig 1b; 26pp; English.
XX
CC      The gene product is known to stimulate cell proliferation and
CC      inhibit anchorage-dependent growth of a variety of human cancer cell
CC      lines, it is esp. useful in treatment of burns and the promotion of
CC      surface and internal wound healing. TGF-beta may be expressed from a
CC      transformed CHO cell line.
CC      (Updated on 31-OCT-2002 to add missing OS field.)
XX
SQ      Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match          73.4%;   Score 990.8;   DB 7;   Length 2537;
Best Local Similarity 85.5%;   Pred. No. 4.5e-193;
Matches 1146;   Conservative    0;   Mismatches 147;   Indels    47;   Gaps        2;

Qy              6  CCGAGATGGCGCCTTCGCGGGCTGCGGGCTCTTGCCGCTGCTGTGCCGCTGCTGTGGCTGC 65
                  ||| ||| |||| | | ||||||||| | ||||||||| | ||||||||| |
Db             837  CCCCCATGCCGCCCTCCGCGGGCTGCGGGCTGCTGCCGCTGCTGTACCGCTGCTGTGGCTAC 896
```

Qy	66	TAGTGTCTGACGCCTGGCCCGGCCGGCCCGCCGGACTGTCCACCCTGCAAGACCATCGACATGG	125
Db	897	TGGTGCTGACGCCTGGCCCCGCCGGCCCGGGACTATCCACCTGCAAGACTATCGACATGG	956
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	957	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	1016
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCTGTCTCG	1076
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACG	1196
Qy	366	AAATCTATGATAAAATTC AAGGGCACCCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	1376
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCAACGGAGTTGTGCGGCAGTGGCTGACCCG CAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCAACGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCC ACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATT CATGGCATGAACCGGCCCTTCCTGC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697

```

Qy      906 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC 965
      |||
Db      1698 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC 1757
      |||

Qy      966 TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT 1025
      |||
Db      1758 TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT 1817
      |||

Qy      1026 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1085
      |||
Db      1818 GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 1877
      |||

Qy      1086 ACAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCCA 1145
      |||
Db      1878 ATAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCCA 1937
      |||

Qy      1146 TCGTGTAATACTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1205
      |||
Db      1938 TCGTGTAATACTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1997
      |||

Qy      1206 CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC 1265
      |||
Db      1998 CCTGCAAGTGCAGCTGAGGTCCC GCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC 2057
      |||

Qy      1266 CCCACCCCCGCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG 1317
      |||
Db      2058 CCCGCCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG 2117
      |||

Qy      1318 CCCACTTGGGATCGATTAAA 1337
      |||
Db      2118 CCCACCTGGGGCCCCATTAA 2137
      |||

```

RESULT 4

AAQ03301

ID AAQ03301 standard; DNA; 2537 BP.

XX

AC AAQ03301;

XX

DT 25-MAR-2003 (updated)

DT 05-AUG-1990 (first entry)

XX

DE cDNA encoding human pre-transforming growth factor-beta-1 (pre-TGF-beta-1).

XX

KW Transforming growth factor-beta-1 (TGF-beta-1);

KW neoplastic cell line inhibition;

KW EGF-potentiates anchorage-independent growth;

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	842..2014
----	-----	-----------

FT		/*tag= a
----	--	----------

FT	mat_peptide	1676..2011
----	-------------	------------

FT		/*tag= b
----	--	----------

FT	misc_difference	37..113
----	-----------------	---------

```

FT          /*tag= c
FT          /note="stable hairpin loops"
FT  misc_feature  2015..2100
FT          /*tag= d
FT          /note="G-C rich sequence
FT          and a downstream TATA-like sequence"
XX
PN  US4886747-A.
XX
PD  12-DEC-1989.
XX
PF  13-MAR-1987;  87US-0025423.
XX
PR  13-MAR-1987;  87US-0025423.
XX
PA  (GETH ) GENENTECH INC.
XX
PI  Derynck RMA,  Goeddel DV;
XX
DR  WPI; 1990-051338/07.
DR  P-PSDB; AAR05258.
XX
PT  Nucleic acid encoding transforming growth factor-beta -
PT  cloned into expression vectors for expression in eukaryotic host
PT  cells for therapeutic use
XX
PS  Disclosure;  Fig 1b; 28pp; English.
XX
CC  It was obtained by an analysis of several overlapping cDNAs and gene
CC  fragments, leading to the detn. of a continuous sequence corresp. to the
CC  TGF-beta-1 precursor mRNA. It is useful in constructing vectors that
CC  encode biologically active transforming growth factor (TGF-beta),
CC  operably linked to DNA that encodes a secretory leader (SL). It, or a
CC  nucleic acid capable of hybridising with it, can also be labelled and
CC  used in diagnostic assays for DNA or mRNA encoding TGF-beta or related
CC  proteins.
CC  (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ  Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match          73.4%;  Score 990.8;  DB 11;  Length 2537;
Best Local Similarity  85.5%;  Pred. No. 4.5e-193;
Matches 1146;  Conservative  0;  Mismatches  147;  Indels  47;  Gaps  2;

Qy          6  CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC  65
   ||  ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          837  CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC  896

Qy          66  TAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG  125
   |  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          897  TGGTGCTGACGCCTGGCCCGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG  956

Qy          126  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC  185
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db          957  AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC  1016

Qy          186  TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG  245

```

Db	1017	TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTACCCCGCTGCTAATGGTGGAAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTACCCCGCTGCTAATGGTGGAAACCCACAACG	1196
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT	1376
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTGATGGCATGAACCGGCCTTCTCTGC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1758	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCT	1817
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACGAGC	1085

```

Db      1818  GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 1877
Qy      1086  ACAACCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA 1145
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1878  ATAACCCGGGCGCCTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCA 1937
Qy      1146  TCGTGTA TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1205
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1938  TCGTGTA TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT 1997
Qy      1206  CCTGCAAGTGCAGCTGAGGCCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC 1265
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1998  CCTGCAAGTGCAGCTGAGGTCCCCGCCCGCCCCGCCCGCCCCGCCCGGGCAGGCCCCGGCCCCAC 2057
Qy      1266  CCCACCCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG 1317
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2058  CCCGCCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG 2117
Qy      1318  CCCACTTGGGATCGATTAAA 1337
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2118  CCCACCTGGGGCCCCATTAA 2137

```

RESULT 5

AAQ02814

ID AAQ02814 standard; cDNA; 2537 BP.

XX

AC AAQ02814;

XX

DT 25-MAR-2003 (updated)

DT 31-OCT-2002 (updated)

DT 31-MAY-1989 (first entry)

XX

DE Sequence of pre-TGF-beta1 cDNA.

XX

KW Transforming growth factor beta-3 (TGF beta 3); tumour cells; growth inhibition.

XX

OS Homo sapiens.

XX

Key	Location/Qualifiers
FT CDS	842..2011
FT	/*tag= a
FT	/label=pre-TGF beta 1
FT CDS	1677..2011
FT	/*tag= b
FT	/label=mature TGF-beta 1
FT GC_signal	2015..2092
FT	/*tag= c
FT misc_feature	2093..2099
FT	/*tag= d
FT	/label=TATA-like sequence
FT stem_loop	37..113
FT	/*tag= e
FT misc_feature	863..911
FT	/*tag= f
FT	/label=hydrophobic domain

Qy	246	CTCTTTTACAACAGTATACCCGCGACCCGGGTAGCCGGGGAAAGTGTCGAACCCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCCGGGTGGCCGGGGAGAGTGTCAGAACCCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCCGCGTGCTAATGGTGGAACCCACAACG	1196
Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCCT	1376
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGAATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGTC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTCTGTC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1758	TGGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1817
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1818	GCCCCTACATTTGGAGCCTGGACACGAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877

[illegible]

RESULT 6

ID AAT15720 standard; cDNA; 2537 BP.

AC

DT 25-MAR-20

DT 25-JAN-1980 (first en

DE Pre-transforming growth factor beta 1 cDNA.

KW transforming growth factor beta 1; wound healing;

XX

XX

FT 5'UTR 1..841

```
FT  misc_feature    37..113
```

```
FT      /note= "GC-rich region forms stable hairpin loops;
FT              similar to structural organisation of c-myc RNA,
FT              could play role in mRNA stability or in
FT              regulation of transcription"
```

FT /*tag= c

FT	mat_peptide	1676..2011	-	-
----	-------------	------------	---	---

FT /product= mature_TGF_beta_1

```
FT          /*tag=  e
```

FT fact 3'UTR of mRNA could not be cloned as cDNA;

FT may be important for transcription efficiency"

FT repeat_unit 2019..2023

FT /*tag= f

FT TATA_signal 2094..2100

FT /*tag= g

FT /note= "TATA-like sequence; no evidence that this

FT functions a promoter"

FT polyA_signal 2514..2520

FT /*tag= h

FT misc_signal 2529..2536

FT /*tag= i

FT /note= "consensus sequence immediately precedes

FT polyA-tail (Benoist et al)"

XX

PN US5482851-A.

XX

PD 09-JAN-1996.

XX

PF 05-NOV-1993; 93US-0147364.

XX

PR 13-MAR-1987; 87US-0025423.

PR 22-MAR-1985; 85US-0715142.

PR 04-AUG-1989; 89US-0389929.

PR 04-MAR-1992; 92US-0845893.

PR 05-NOV-1993; 93US-0147364.

XX

PA (GETH) GENENTECH INC.

XX

PI Derynck RMA, Goeddel DV;

XX

DR WPI; 1996-076891/08.

DR P-PSDB; AAR90827.

XX

PT New recombinant human transforming growth factor-beta prods. - produced

PT using Chinese hamster ovary cells, for use in diagnostic applications

PT or in therapy

XX

PS Example 3; Fig 1; 26pp; English.

XX

CC The cDNA encodes the pre-transforming growth factor (TGF) beta 1 protein.

CC The nucleotide sequence was obt'd. by an analysis of several overlapping

CC cDNAs and gene fragments. The DNA is useful for the recombinant

CC production of TGF beta 1, which can be used in, e.g. wound healing.

CC (Revised entry submitted to correct sequence analysis breakdown.)

CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 2537 BP; 473 A; 893 C; 739 G; 432 T; 0 other;

Query Match 73.4%; Score 990.8; DB 17; Length 2537;

Best Local Similarity 85.5%; Pred. No. 4.5e-193;

Matches 1146; Conservative 0; Mismatches 147; Indels 47; Gaps 2;

QY 6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65

|| ||| |||| || ||||||||| ||||||||||| ||||||||||| ||

Db 837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

QY 66 TAGTGCTGACGCCTGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGG 125

Db	897	<div> </div> TGGTGTGACGCCTGGCCCCGCCGGCCGCGGGACTATCCACCTGCAAGACTATCGACATGG	956
Qy	126	<div> </div> AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	957	<div> </div> AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	1016
Qy	186	<div> </div> TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG	245
Db	1017	<div> </div> TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG	1076
Qy	246	<div> </div> CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC	305
Db	1077	<div> </div> CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	306	<div> </div> CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	1137	<div> </div> CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACG	1196
Qy	366	<div> </div> AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	<div> </div> AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	<div> </div> AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	<div> </div> AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	<div> </div> GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT	545
Db	1317	<div> </div> GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT	1376
Qy	546	<div> </div> GGCCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1377	<div> </div> GGCATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	<div> </div> ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCGAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1437	<div> </div> ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	<div> </div> TCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	<div> </div> TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	<div> </div> TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTGTC	785
Db	1557	<div> </div> TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTGTC	1616
Qy	786	<div> </div> TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	<div> </div> TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	<div> </div> CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1676	<div> </div> -----GCCCTGGACACCAACTATTGCT	1697
Qy	906	<div> </div> TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965

Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	966	TGGGCTGGAAGTGGATTATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1758	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCT	1817
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1818	GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877
Qy	1086	ACAACCCGGGCGCGTCTGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1145
Db	1878	ATAACCCGGGCGCCTCTGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA	1937
Qy	1146	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1205
Db	1938	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1997
Qy	1206	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCCCACCCGGCAGGCCCCGGC	1265
Db	1998	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCCGCGCAGGCCCCGGCCCCAC	2057
Qy	1266	CCCACCCCCGCCCCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1317
Db	2058	CCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG	2117
Qy	1318	CCCACCTGGGATCGATTAAA	1337
Db	2118	CCCACCTGGGGCCCCATTAA	2137

RESULT 7

AAQ56923

ID AAQ56923 standard; cDNA; 2537 BP.

XX

AC AAQ56923;

XX

DT 25-MAR-2003 (updated)

DT 09-JUL-1994 (first entry)

XX

DE Human pre-TGF-beta-1.

XX

KW TGF-beta-1; TGF-beta-2; transforming growth factor beta-1;

KW transforming growth factor beta-3; recombinant; wound healing;

KW vulnerary; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT misc structure 47..113

```
FT                               /*tag=  a
```

```
FT      /note= "possible hairpin loop region"
```

FT CDS 842..2014

FT /*tag= b

FT mat_peptide 1676..2011

FT /*tag= c

FT polyA signal 2515..2521

Qy	246	CTCTTTACAACAGTACCCGCGACCCGGGTAGCCGGGGAAAAGTGTGCGAACCCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCCGGGTGGCCGGGGAGAGTGCAGAACCCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACG	1196
Qy	366	AAATCTATGATAAAATTCAGGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAGGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1376
Qy	546	GGCGCTACCTCAGCAACCCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1377	GGCGATACCTCAGCAACCCGGCTGCTGGCACCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1616
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1758	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT	1817
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1818	GCCCCACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877
Qy	1086	ACAACCCGGGCGCGTGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1145

Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGC	305
Db	1077	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC	1136
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	1137	CTGAGGCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACG	1196
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1197	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	1256
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT	545
Db	1317	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT	1376
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1377	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1436
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1437	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGAGAAATTGAGGGCTTTTCGCC	1496
Qy	666	TCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1497	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1556
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1557	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCTCTGC	1616
Qy	786	TCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1617	TTCTCATGGCCACCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1675
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1676	-----GCCCTGGACACCAACTATTGCT	1697
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1698	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1757
Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1758	TCGGCTGGAAGTGGATCCACGAGCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT	1817
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1818	GCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1877
Qy	1086	ACAACCCGGGCGCGTGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA	1145

Db	1878	ATAAACCCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCG CAGGCGCTGGAGCCGCTGCCCA	1937
Qy	1146	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1205
Db	1938	TCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1997
Qy	1206	CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCCGGC	1265
Db	1998	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGCCGAGGCCCGGCCAC	2057
Qy	1266	CCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1317
Db	2058	CCCGCCCCGCCCGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAAG	2117
Qy	1318	CCCACCTGGGATCGATTAAA	1337
Db	2118	CCCACCTGGGGCCCCATTAA	2137

RESULT 9

AAQ03268

ID AAQ03268 standard; DNA; 1561 BP.

XX

AC AAQ03268;

XX

DT 25-MAR-2003 (updated)

DT 12-AUG-1990 (first entry)

XX

DE Simian transforming growth factor-beta cDNA.

XX

KW Transforming growth factor-beta; psoriasis; TGF-beta; ss.

XX

OS Monkey.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	sig peptide	283..324
----	-------------	----------

```
FT      /*tag=  a
```

FT mat_peptide 1096..1431

FT /*tag= b

FT /product=human transforming growth factor-beta

XX

PN EP353772-A.

XX

PD 07-FEB-1990.

XX

PF 04-AUG-1989; 89EP-0114458.

XX

PR 05-AUG-1988; 88US-0229133.

XX

PA (ONCO) ONCOGEN LP.

XX

PI Twardzik DR, Purchio AF, Ranchalis JE, Stevens V;

XX

DR WPI; 1990-038499/06.

DR P-PSDB; AAR03743.

XX

PT Inhibition of proliferation of epidermal cells -
PT used to treat psoriasis by contacting cells with compositions
PT containing transforming growth factor-beta.

XX

PS Disclosure; fig 1; 20pp; English.

XX

CC TGF-beta may be used in the treatment of hyperplasia
CC associated with acanthosis-categorised skin diseases, and
CC in alleviating psoriatic symptoms associated with cytokine-
CC induced phenomena. See also AAQ03269 and AAR03750.
CC (Updated on 25-MAR-2003 to correct PA field.)

XX

SQ Sequence 1561 BP; 301 A; 547 C; 446 G; 267 T; 0 other;

Query Match 72.9%; Score 983; DB 11; Length 1561;
Best Local Similarity 85.6%; Pred. No. 1.6e-191;
Matches 1148; Conservative 0; Mismatches 145; Indels 48; Gaps 3;

```
Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    257 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 316

Qy     66 TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    317 TGGTGCTGACGCCTAGCCGGCCGGCCGGCAGGACTATCCACCTGCAAGACTATCGACATGG 376

Qy    126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    377 AGCTGGTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 436

Qy    186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    437 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 496

Qy    246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    497 CCCTGTACAACAGCACCCGCGACCGGGTGCCGGGGAGAGTGCGGAGCCGGAGCCCGAAC 556

Qy    306 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 365
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    557 CGGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG 616

Qy    366 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    617 AAATCTATGACAAGTTCAAGCAGAGCACACAGCATATATATGTTCTTCAACACATCAG 676

Qy    426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    677 AGCTCCGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 736

Qy    486 GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    737 GGCTCAAGTTAAAAGTGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCCT 796

Qy    546 GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCCTTTG 605
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    797 GGCGATACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTTG 856
```


XX
 AC AAI58342;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 545.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR P-PSDB; AAM39186.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 545; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX

SQ Sequence 2742 BP; 526 A; 938 C; 799 G; 479 T; 0 other;

Query Match 72.8%; Score 981.4; DB 22; Length 2742;
Best Local Similarity 85.5%; Pred. No. 3.8e-191;
Matches 1147; Conservative 0; Mismatches 146; Indels 48; Gaps 3;

```
Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      ||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy     66 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    897 TGGTGCTGACGCCTGGCCGGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy    126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy    186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCGGCCCGCTGCCTGAGGCAGTACTGG 245
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCGGCCCGCTGCCCGAGGCCGTGCTCG 1076

Qy    246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy    306 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 365
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1137 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG 1196

Qy    366 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy    426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1257 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 1316

Qy    486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTTCCT 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1317 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTTCCT 1376

Qy    546 GGCCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGCGTGTCTTTTG 605
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1377 GGCATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTATCTTTTG 1436

Qy    606 ATGTCACCGGAGTTGTGCGGCAGTGCGTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC 665
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1437 ATGTCACCGGAGTTGTGCGGCAGTGCGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC 1496

Qy    666 TCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT 725
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

Db      1497 TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT 1556
Qy      726 TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTTCACGGCATGAACCGGCCCTTCCTGC 785
      ||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1557 TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTTCATGGCATGAACCGGCCTTTCTCTGC 1616
Qy      786 TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG 845
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1617 TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA- 1675
Qy      846 CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT 905
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1676 -----GCCCTGGACACCAACTATTGCT 1697
Qy      906 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC 965
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1698 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGGAAGGACC 1757
Qy      966 TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT 1025
      | ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| |||||
Db     1758 TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT 1817
Qy     1026 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1085
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1818 GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 1877
Qy     1086 ACAACCCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA 1145
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1878 ATAACCCGGGCGCCTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCA 1937
Qy     1146 TCGTGTAATACTGGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1205
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1938 TCGTGTAATACTGGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT 1997
Qy     1206 CCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCCCACAGCCCCGCCCCACCCGGCAGGCCCCGGC 1265
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     1998 CCTGCAAGTGCAGCTGAGGTCCC GCCCCCCGCCCCGCCCCGCCCCGGCAGGCCCCGGCCCCAC 2057
Qy     1266 CCCACCCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAA 1316
      ||| |||| ||| || || ||||| ||||| ||||| ||||| |||||
Db     2058 CCCGCCCCGCCCCGCTGCTTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCAA 2117
Qy     1317 GCCCACTTGGGATCGATTAAA 1337
      ||||| |||| | |||
Db     2118 GCCCACCTGGGGCCCCATTAA 2138

```

RESULT 11

AAQ20289

ID AAQ20289 standard; cDNA; 1559 BP.

XX

AC AAQ20289;

XX

DT 25-MAR-2003 (updated)

DT 16-APR-1992 (first entry)

XX

DE Sequence encoding simian transforming growth factor (TGF) beta-1.

```

XX
KW Hypertension therapy; hypotensive agent; blood pressure modulator;
KW ss.
XX
OS Monkey.
XX
FH Key Location/Qualifiers
FT CDS 262..282
FT /*tag= a
FT sig_peptide 283..324
FT /*tag= b
FT CDS 325..1098
FT /*tag= c
FT mat_peptide 1099..1436
FT /*tag= d
XX
PN WO9119513-A.
XX
PD 26-DEC-1991.
XX
PF 20-JUN-1991; 91WO-US04449.
XX
PR 20-JUN-1990; 90US-0541221.
XX
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI Oleson FB, Comereski CR;
XX
DR WPI; 1992-024199/03.
DR P-PSDB; AAR20124.
XX
PT Use of transforming growth factor (TGF)-beta and their
PT antagonists - for modulating blood pressure, for treating
PT hypertension and hypotension
XX
PS Disclosure; Fig 1; 42pp; English.
XX
CC A new method for treating hypertension comprises administering a
CC transforming growth factor (TGF)-beta to an individual at a dose
CC effective for lowering blood pressure; the TGF-beta may be e.g.
CC mature TGF-beta, TGF-beta2, a mature TGF-betal/beta2 hybrid, TGF-
CC betal precursor, a latent TGF-beta2 precursor, hybrid TGF-betal/TGF-
CC beta2 precursor, a latent TGF-betal complex or a latent TGF-beta2
CC complex.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1559 BP; 300 A; 546 C; 446 G; 267 T; 0 other;

Query Match 72.7%; Score 980.6; DB 13; Length 1559;
Best Local Similarity 85.7%; Pred. No. 5.1e-191;
Matches 1144; Conservative 0; Mismatches 144; Indels 47; Gaps 3;

Qy 11 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 70
   ||| |||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 320

Qy 71 CTGACGCCTGGCCGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 130

```

Db	321	CTGACGCCTAGCCGGCCGGCCGAGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	380
Qy	131	GTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	190
Db	381	GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	440
Qy	191	AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT	250
Db	441	AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTG	500
Qy	251	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG	310
Db	501	TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCGAGCCGAACCGGAG	560
Qy	311	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	370
Db	561	GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC	620
Qy	371	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	430
Db	621	TATGACAAGTTCAAGCAGAGCACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	431	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	490
Db	681	CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	491	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC	550
Db	741	AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGA	800
Qy	551	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGAAGTCAACCGGAGTGGCTGTCTTTGATGTC	610
Db	801	TACCTCAGCAACCGGCTGCTGGCGCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	611	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGT	670
Db	861	ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCTTAGC	920
Qy	671	GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT	730
Db	921	GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT	980
Qy	731	TCTGGCCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	790
Db	981	ACCGGCCCGCCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCCTGCTTCTC	1040
Qy	791	ATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	850
Db	1041	ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-----	1094
Qy	851	GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC	910
Db	1095	-----GCCCTGGACACCAACTACTGCTTCAGC	1121
Qy	911	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC	970

```

Db      1122 TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGC 1181
Qy      971 TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC 1030
        |||||
Db      1182 TGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTGGGGCCCTGTCCC 1241
Qy      1031 TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCTCTGTACAACCAGCACAAAC 1090
        |||||
Db      1242 TACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTCTGGCCCTGTACAACCAGCATAAC 1301
Qy      1091 CCGGGCGCGTTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG 1150
        |||||
Db      1302 CCGGGCGCCTCGGCGGCGCGGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG 1361
Qy      1151 TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC 1210
        |||||
Db      1362 TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC 1421
Qy      1211 AAGTGCAGCTGAGGCCCCGCCCCGCCCCACAGCCCCGCCCCACCCGGCAGGCCCCGGCCCCAC 1270
        || |||||
Db      1422 AAATGCAGCTGAGGCCCCGCCCCGCCCCGCCCCACCCCGGCAGGCGCGCCCCGCCCCACC 1481
Qy      1271 CCCCCCGCGC-----CTCACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAAGCCCAC 1322
        || |||||
Db      1482 CCACCCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTGCCCCAAGCCCAC 1541
Qy      1323 TTGGGATCGATTAAA 1337
        |||||
Db      1542 CTGGGGCCCCATTAA 1556

```

RESULT 12

AAQ13392

ID AAQ13392 standard; DNA; 1821 BP.

XX

AC AAQ13392;

XX

DT 20-NOV-1991 (first entry)

XX

DE Human pro-TGF-beta 1 gene.

XX

KW Osteogenetic; tumoricidal; ss.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	512..1684
----	-----	-----------

FT		/*tag= a
----	--	----------

FT	sig_peptide	512..598
----	-------------	----------

FT		/*tag= b
----	--	----------

FT	misc_RNA	599..1684
----	----------	-----------

FT		/*tag= c
----	--	----------

FT		/note= "pro-TGF-beta 1"
----	--	-------------------------

FT	mat_peptide	1346..1684
----	-------------	------------

FT		/*tag= e
----	--	----------

FT		/note= "TGF-beta 1"
----	--	---------------------

XX

Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTTGCTGA	485
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	986
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTCT	545
Db	987	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTCT	1046
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGTGTCTTTG	605
Db	1047	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTATCTTTTG	1106
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCC	665
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1166
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1167	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGACATCAACGGGT	1226
Qy	726	TCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	785
Db	1227	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTTCATGGCATGAACCGGCCCTTCCTGC	1286
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1346	-----GCCCTGGACACCAACTATTGCT	1367
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1368	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1427
Qy	966	TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1428	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT	1487
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1488	GCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1547
Qy	1086	ACAACCCGGGCGCGTGGCGGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA	1145
Db	1548	ATAACCCGGGCGCCTCGGCGGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA	1607
Qy	1146	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1205
Db	1608	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1667
Qy	1206	CCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC	1265
Db	1668	CCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCCAGGCCCGGCCAC	1727
Qy	1266	CCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAG	1317

```

      ||| |||| ||| ||      || |||||||||||||||| ||||||| |
Db      1728 CCCGCCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCCCAA 1787
QY      1318 CCC 1320
      ||
Db      1788 GCC 1790

```

RESULT 13

AAN81084

ID AAN81084 standard; cDNA; 1560 BP.

XX

AC AAN81084;

XX

DT 25-MAR-2003 (updated)

DT 09-OCT-1990 (first entry)

XX

DE Coding sequence of simian transforming growth factor-beta 1.

XX

KW Transforming growth factor-beta 1; tumour treatment; ss cDNA.

XX

OS Cercopithecus aethiops.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	261..1433
----	-----	-----------

FT		/*tag= a
----	--	----------

FT	sig_peptide	282..323
----	-------------	----------

FT		/*tag= b
----	--	----------

FT	mat_peptide	1095..1433
----	-------------	------------

FT		/*tag= c
----	--	----------

XX

PN EP293785-A.

XX

PD 07-DEC-1988.

XX

PF 27-MAY-1988; 88EP-0108528.

XX

PR 29-MAY-1987; 87US-0055662.

PR 25-JAN-1988; 88US-0147842.

XX

PA (ONCO) ONCOGEN.

PA (BRIM) BRISTOL-MYERS CO.

XX

PI Purchio AG, Gentry L, Twardzik D;

XX

DR WPI; 1988-347488/49.

DR P-PSDB; AAP80647.

XX

PT Prodn. of simian transforming growth factor beta-1 - by culturing transfected eucaryotic cells, and new precursor proteins, useful for treating tumours.

XX

PS Disclosure; Page ?; pp; English.

XX

CC The cDNA is prepd. from African green monkey cell line BSC-40 and is expressed in eukaryotic cells in plasmid pSV2. There is 100% homology between mature simian and human TGF-beta 1. The plasmid also contains

CC the SV40 promoter and a selection marker, esp. DHFR.
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)

XX

SQ Sequence 1560 BP; 301 A; 547 C; 445 G; 267 T; 0 other;

Query Match 72.6%; Score 979.6; DB 9; Length 1560;
Best Local Similarity 85.6%; Pred. No. 8.2e-191;
Matches 1144; Conservative 0; Mismatches 144; Indels 48; Gaps 3;

```
Qy      11 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 70
      ||| |||| || ||||||||||| ||||||||||| ||||||||||| ||| |||
Db     261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 320

Qy      71 CTGACGCCTGGCCGGCCGGCCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 130
      ||||||||| ||||||||||| ||||| ||||||||||| ||||||||| |||||
Db     321 CTGACGCCTAGCCGGCCGGCCCGCAGGACTATCCACCTGCAAGACTATCGACATCGAGCTG 380

Qy     131 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 190
      ||||||||||| ||||||||| ||||| ||||||||||| ||||||||| ||||| |||
Db     381 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 440

Qy     191 AGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTT 250
      ||||||||| ||||||||| ||||||||| ||||||||| ||||| ||| ||| |||
Db     441 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCCGAGGCCGTGCTCGCCCTG 500

Qy     251 TACAACAGTACCCGCGACCGGGTAGCCGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAG 310
      ||||||||| ||||||||||| ||||||||| ||||| || ||||||||| ||| |||
Db     501 TACAACAGCACCCGCGACCGGGTGGCCGGGAGAGTGCAGGAGCCGAGCCCGAACCGGAG 560

Qy     311 GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACCAAATC 370
      || ||||||||||| ||||||||||| ||||||||||| ||||| || ||||| |||||
Db     561 GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACGAAATC 620

Qy     371 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 430
      ||||| || ||||||| ||| || ||||||| ||||||||| ||||||||| || |||||
Db     621 TATGACAAGTTCAAGCAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC 680

Qy     431 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 490
      || |||||| || || ||||||||| ||||||||| ||||||||| ||||||||| |||
Db     681 CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGAGGCTC 740

Qy     491 AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGC 550
      ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||
Db     741 AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA 800

Qy     551 TACCTCAGCAACCGGCTGCTGGCCCCCAGTGAAGTCAACCGGAGTGGCTGTCTTTGATGTC 610
      ||||||||||| ||||||||| ||||| ||||| ||||||||| ||||| |||||||||
Db     801 TACCTCAGCAACCGGCTGCTGGCGCCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC 860

Qy     611 ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCCTCAGT 670
      ||||||||||| ||||||||| ||| ||||| ||| || ||||| ||||||||| |||
Db     861 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCCCTTAGC 920

Qy     671 GCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAAT 730
      ||||||||| ||||| ||||||||||| ||||||||| || ||||| || |||||||||
Db     921 GCCCACTGCTCCTGTGACAGCAAAGATAACACACTGCAAGTGGACATCAACGGGTTCACT 980
```


Qy	731	TCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTC	790
Db	981	ACCGGCCGCCGAGGTGACCTGGCCACAATTCATGGCATGAACCGGCCCTTCCTGCTTCTC	1040
Qy	791	ATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTG	850
Db	1041	ATGGCCACCCCGCTGGAGAGGGCCCAACATCTGCAAAGCTCCCGGCACCGCCGA-----	1094
Qy	851	GATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGC	910
Db	1095	-----GCCCTGGACACCAACTACTGCTTCAGC	1121
Qy	911	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGC	970
Db	1122	TCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTATATTGACTTCCGCAAGGACCTCGGC	1181
Qy	971	TGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1030
Db	1182	TGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTGGGGCCCTGTCCC	1241
Qy	1031	TACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTTGCTCTGTACAACCAGCACAAAC	1090
Db	1242	TACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGCCCTGTACAACCAGCATAAC	1301
Qy	1091	CCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1150
Db	1302	CCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG	1361
Qy	1151	TACTACGTGGGCCCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC	1210
Db	1362	TACTACGTGGGCCCGCAAGCCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGA	1421
Qy	1211	AAGTGCAGCTGAGGCCCCGCCCCGCCCCACAGCCCCGCCCCACCGGCAGGCCCCGGCCCCAC	1270
Db	1422	AAATGCAGCTGAGGCCCCGCCCCGCCCCGCCCCACCGGCAGGCCCCGGCCCCGCCCCAC	1481
Qy	1271	CCCCGCCCCGCT-----CACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAAGCCCCA	1321
Db	1482	CCCACCCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTGCCCCAAGCCCCA	1541
Qy	1322	CTTGGGATCGATTAAA	1337
Db	1542	CCTGGGGCCCCATTAA	1557

RESULT 14

AAQ03508

ID AAQ03508 standard; DNA; 1560 BP.

XX

AC AAQ03508;

XX

DT 25-MAR-2003 (updated)

DT 09-JAN-2003 (updated)

DT 14-AUG-1990 (first entry)

XX

DE Simian Transforming growth factor - Beta1.

XX
 KW HIV; AIDS; SIV; vaccine; AZT; CD4; cytokines; growth
 KW factors; ds.
 XX
 OS Cebus apella.
 XX
 FH Key Location/Qualifiers
 FT CDS 267..1437
 FT /*tag= a
 FT mat_peptide 1103..1437
 FT /*tag= b
 XX
 PN EP356935-A.
 XX
 PD 07-MAR-1990.
 XX
 PF 25-AUG-1989; 89EP-0115719.
 XX
 PR 25-AUG-1988; 88US-0236698.
 XX
 PA (ONCO) ONCOGEN LP.
 XX
 PI Brankovan V, Lioubin M, Purchio A;
 XX
 DR WPI; 1990-068723/10.
 DR P-PSDB; AAR05663.
 XX
 PT Compsns. contg. transforming growth factor beta -
 PT used for inhibitions of HIV infection and replication in vivo.
 XX
 PS Disclosure; Fig 1; 20pp; English.
 XX
 CC TGF-beta may be used in vivo to prevent formation of syncytia and
 CC inhibit HIV infection. TGF may also be used with other HIV treatments
 CC (AZT, soluble CD4 etc.).
 CC (Updated on 09-JAN-2003 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 1560 BP; 301 A; 547 C; 445 G; 267 T; 0 other;

Query Match 72.6%; Score 979.6; DB 11; Length 1560;
 Best Local Similarity 85.6%; Pred. No. 8.2e-191;
 Matches 1144; Conservative 0; Mismatches 144; Indels 48; Gaps 3;

Qy 11 ATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG 70
 ||| |||| || ||||||||| ||||||||||| ||||||||||| || |||
 Db 261 ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG 320
 Qy 71 CTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG 130
 ||||||||| ||||||||||| ||||| ||||||||||| ||||||||| |||||
 Db 321 CTGACGCCTAGCCGGCCGGCCGAGGACTATCCACCTGCAAGACTATCGACATCGAGCTG 380
 Qy 131 GTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCC 190
 ||||||||||| ||||||||||| ||||| ||||||||||| ||||||||| |||||
 Db 381 GTGAAGCGGAAGCGCATCGAGACCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC 440
 Qy 191 AGCCCCCGAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTT 250

Db	441	 AGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGTGCCCGAGGCCGTGCTCGCCCTG	500
Qy	251	TACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCGAGCCAGAG	310
Db	501	 TACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGGAGCCCGAGCCCGAACCGGAG	560
Qy	311	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATC	370
Db	561	 GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATC	620
Qy	371	TATGATAAAATTCAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC	430
Db	621	 TATGACAAGTTCAGGAGAGCACACACAGCATATATATGTTCTTCAACACATCAGAGCTC	680
Qy	431	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC	490
Db	681	 CGAGAAGCAGTACCTGAACCTGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTC	740
Qy	491	AAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGC	550
Db	741	 AAGTTAAAAGTCGAGCAGCATGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGA	800
Qy	551	TACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTTTGATGTC	610
Db	801	 TACCTCAGCAACCGGCTGCTGGCGCCCAGCAACTCGCCGGAGTGGTTGTCTTTGATGTC	860
Qy	611	ACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGT	670
Db	861	 ACCGGAGTTGTGCGGCAGTGGTTGAGCCGCGGAGGGGAAATTGAGGGCTTTTCGCCTTAGC	920
Qy	671	GCCCACTCTTCCTCTGACAGCAAAGATAACA CACTCCACGTGGAAATTAACGGGTTCAAT	730
Db	921	 GCCCACTGCTCCTGTGACAGCAAAGATAACA CACTGCAAGTGGACATCAACGGGTTCACT	980
Qy	731	TCTGGCCCGCCGGGGTGACCTGGCCACCA TTACGGCATGAACCGGCCCTTCCTGCTCCTC	790
Db	981	 ACCGGCCCGCCGAGGTGACCTGGCCACAATT CATGGCATGAACCGGCCCTTCCTGCTTCTC	1040
Qy	791	ATGGCCACCCCGCTGGAGAGGGCCCAGCAC CTGCACAGCTCCCGGCACCGCCGAGCCCTG	850
Db	1041	 ATGGCCACCCCGCTGGAGAGGGCCCAACAT CTGCAAAGCTCCCGGCACCGCCGA-----	1094
Qy	851	GATACCAACGACTACAAGGATGACGACGACA AGGCCCTGGATACCAACTACTGCTTCAGC	910
Db	1095	 -----GCCCTGGACACCAACTACTGCTTCAGC	1121
Qy	911	TCCACGGAGAAGAACTGCTGCGTGCGGCAG CTCTACATTGACTTCCGGAAGGACCTGGGC	970
Db	1122	 TCCACGGAGAAGAACTGCTGCGTGCGGCAG CTGTATATTGACTTCCGCAAGGACCTCGGC	1181
Qy	971	TGGAAGTGGATTTCATGAACCCAAGGGCTA CCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1030
Db	1182	 TGGAAGTGGATCCACGAGCCCAAGGGCTA CCATGCCAATTTCTGCCTGGGGCCCTGTCCC	1241
Qy	1031	TACATCTGGAGCCTAGACACTCAGTACAGCA AGGTCTGGCTCTGTACAACCAGCACAAAC	1090

```

Db      1242 TACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTTGGCCCTGTACAACCAGCATAAC 1301
Qy      1091 CCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG 1150
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1302 CCGGGCGCCTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTG 1361
Qy      1151 TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGC 1210
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1362 TACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGA 1421
Qy      1211 AAGTGCAGCTGAGGCCCGCCCCGCCCCCAGCCCCGCCCCACCCGGCAGGCCCGGCCCCAC 1270
          || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1422 AAATGCAGCTGAGGCCCGCCCCGCCCCCAGCCCCACCCGGCAGGCCCGGCCCGGCCCCAC 1481
Qy      1271 CCCC GCCCGCCT-----CACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAAGCCCA 1321
          ||| ||| || | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1482 CCCACCCCGCTGTCTTGCCCTTGGGGGCTGTATTTAAGGACACCCGTGCCCCAAGCCCA 1541
Qy      1322 CTTGGGATCGATTAAA 1337
          ||||| | |||
Db      1542 CCTGGGGCCCCATTAA 1557

```

RESULT 15

AAT05876

ID AAT05876 standard; cDNA; 2745 BP.

XX

AC AAT05876;

XX

DT 25-JUN-1996 (first entry)

XX

DE cDNA encoding transforming growth factor-beta 1.

XX

KW macrophage inducible nitric oxide synthase; iNOS; constitutive NOS;

KW interleukin-1-beta; transforming growth factor-beta; TGF-beta; IL1-beta;

KW nitric oxide production; hypotension; inflammation; septic shock;

KW treatment; ds.

XX

OS Mammalian sp.

XX

FH Key Location/Qualifiers

FT CDS 842..2017

FT /*tag= a

FT /product= transforming growth factor-beta 1

XX

PN WO9526745-A1.

XX

PD 12-OCT-1995.

XX

PF 05-APR-1994; 94WO-US03705.

XX

PR 05-APR-1994; 94WO-US03705.

XX

PA (HARD) HARVARD COLLEGE.

XX

PI Lee M, Perrella MA;

XX

[illegible]

Db	1317	GGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATT	1376
Qy	543	CCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTACCCGGAGTGGCTGTCTT	602
Db	1377	CCTGGCGATACTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTT	1436
Qy	603	TTGATGTACACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTC	662
Db	1437	TTGATGTACACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTC	1496
Qy	663	GCCTCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACG	722
Db	1497	GCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACG	1556
Qy	723	GGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCC	782
Db	1557	GGTTCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGCGCATGAACCGGCCCTTTC	1616
Qy	783	TGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCC	842
Db	1617	TGCTTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCC	1676
Qy	843	GAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACT	902
Db	1677	GA-----GCCCTGGACACCAACTATT	1697
Qy	903	GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGG	962
Db	1698	GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGG	1757
Qy	963	ACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGC	1022
Db	1758	ACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGC	1817
Qy	1023	CCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTCTGGCTCTGTACAACC	1082
Db	1818	CCTGCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTCTGGCCCTGTACAACC	1877
Qy	1083	AGCACAACCCGGGCGCGTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGC	1142
Db	1878	AGCATAAACCCGGGCGCCTTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGC	1937
Qy	1143	CCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1202
Db	1938	CCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1997
Qy	1203	GTTCTCTGCAAGTGCAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCC	1262
Db	1998	GCTCCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCAGGCCCGGCC	2057
Qy	1263	GGCCCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCC	1314
Db	2058	CACCCCGCCCCCGCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCC	2117
Qy	1315	AAGCCCACTTGGGATCGATTAAA	1337
Db	2118	AAGCCCACTTGGGGCCCCATTAA	2140

Search completed: October 27, 2003, 19:11:13
Job time : 415.282 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2003, 18:35:27 ; Search time 400.963 Seconds
(without alignments)
9022.658 Million cell updates/sec

Title: US-10-017-372E-32
Perfect score: 1349
Sequence: 1 tggtagcgagatggcgccctt.....cgattaaagcggccgcgact 1349

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	% Query		Length	DB	ID	Description	
	No.	Score					
	1	981.8	72.8	1821	14	US-10-087-268-1	Sequence 1, Appli
	2	981.4	72.8	2742	14	US-10-037-270-220	Sequence 220, App
	3	980.2	72.7	1821	14	US-10-087-268-4	Sequence 4, Appli
	4	977.8	72.5	2745	11	US-09-948-002-28	Sequence 28, Appl
	5	859	63.7	2094	11	US-09-948-002-1	Sequence 1, Appli
	6	839.4	62.2	1585	11	US-09-948-002-27	Sequence 27, Appl
	7	654	48.5	1376	10	US-09-756-283A-19	Sequence 19, Appl
	8	591.4	43.8	1352	10	US-09-756-283A-21	Sequence 21, Appl
	9	366.4	27.2	489	11	US-09-911-904-167	Sequence 167, App
	10	308.6	22.9	339	10	US-09-813-271B-1	Sequence 1, Appli
	11	238.4	17.7	2574	11	US-09-906-158-3	Sequence 3, Appli
	12	238.4	17.7	2574	13	US-10-028-158-20	Sequence 20, Appl
	13	233.2	17.3	4382	12	US-09-957-458B-9	Sequence 9, Appli
	14	230	17.0	2879	11	US-09-906-158-10	Sequence 10, Appl
	15	227.2	16.8	336	10	US-09-813-271B-7	Sequence 7, Appli
	16	209.4	15.5	339	10	US-09-813-271B-5	Sequence 5, Appli
c	17	185.4	13.7	2381	12	US-10-311-455-2384	Sequence 2384, Ap
	18	184.8	13.7	4267	11	US-09-948-002-47	Sequence 47, Appl
	19	182.4	13.5	336	10	US-09-813-271B-11	Sequence 11, Appl
	20	181.2	13.4	2381	12	US-10-311-455-2383	Sequence 2383, Ap
	21	180.8	13.4	336	10	US-09-813-271B-9	Sequence 9, Appli
	22	160.4	11.9	2570	12	US-09-960-706-663	Sequence 663, App
	23	160.4	11.9	2912	13	US-10-044-090-323	Sequence 323, App
	24	158.2	11.7	339	10	US-09-813-271B-3	Sequence 3, Appli
c	25	141.6	10.5	597	9	US-09-864-761-15319	Sequence 15319, A
c	26	134.8	10.0	154	9	US-09-864-761-31841	Sequence 31841, A
c	27	131.8	9.8	206	10	US-09-833-381-577	Sequence 577, App
c	28	123.2	9.1	537	9	US-09-864-761-8844	Sequence 8844, Ap
c	29	122.2	9.1	148	9	US-09-864-761-25510	Sequence 25510, A
	30	113.8	8.4	181	10	US-09-833-381-1416	Sequence 1416, Ap
	31	104.8	7.8	851	13	US-10-027-632-152938	Sequence 152938,
	32	102.2	7.6	29000	11	US-09-906-158-17	Sequence 17, Appl
c	33	100.4	7.4	224	10	US-09-833-381-73	Sequence 73, Appl
	34	79.4	5.9	450	11	US-09-918-995-6145	Sequence 6145, Ap
c	35	69.8	5.2	431	9	US-09-864-761-18644	Sequence 18644, A
c	36	69.4	5.1	176	10	US-09-833-381-571	Sequence 571, App
	37	69	5.1	584	12	US-10-029-386-9758	Sequence 9758, Ap
	38	68.2	5.1	176	12	US-10-029-386-23458	Sequence 23458, A
c	39	64	4.7	363	10	US-09-833-381-585	Sequence 585, App
	40	62.6	4.6	364	12	US-10-029-386-26476	Sequence 26476, A
	41	62.6	4.6	544	12	US-10-029-386-12776	Sequence 12776, A
	42	62.6	4.6	927	12	US-10-244-718-1	Sequence 1, Appli
	43	57.4	4.3	486	11	US-09-918-995-25641	Sequence 25641, A
	44	55.2	4.1	594	13	US-10-027-632-141376	Sequence 141376,
	45	53.2	3.9	658	9	US-09-765-527-254	Sequence 254, App

ALIGNMENTS

RESULT 1
US-10-087-268-1
; Sequence 1, Application US/10087268


```

; Publication No. US20030119010A1
; GENERAL INFORMATION:
; APPLICANT: Jonsonn, Julie Ruth
; APPLICANT: Powell, Elizabeth Ellen
; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or
condition
; FILE REFERENCE: Fibrosis
; CURRENT APPLICATION NUMBER: US/10/087,268
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1821
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(511)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (512)..(1684)
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (512)..(598)
; OTHER INFORMATION:
; NAME/KEY: 3'UTR
; LOCATION: (1685)..(1821)
; OTHER INFORMATION:
US-10-087-268-1

```

```

Query Match          72.8%; Score 981.8; DB 14; Length 1821;
Best Local Similarity 85.7%; Pred. No. 3.6e-254;
Matches 1134; Conservative 0; Mismatches 142; Indels 47; Gaps 2;

```

```

Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      || ||| |||| || ||||||||||| ||||||||||| ||||||||||| ||
Db      507 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 566

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||||||||||||||||||||| ||||| ||||||||| ||||||||| ||
Db      567 TGGTGCTGACGCCTGGCCGGCCGGCCGGCGGACTATCCACCTGCAAGACTATCGACATGG 626

Qy      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||||||||||||||||||||||| ||||||||||| ||||||||| ||||| |||
Db      627 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 686

Qy      186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||||||||||||||||||||| ||||||||||| ||||||||| ||||| || ||
Db      687 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 746

Qy      246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      | || ||||||||| ||||||||||| ||||||||| ||||| ||||||||| ||
Db      747 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 806

Qy      306 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC 365
      | ||||| ||||||||||| ||||||||||| ||||||||||| || |||||
Db      807 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACG 866

```

Qy	366	AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCCG	425
Db	867	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	926
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	986
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCT	545
Db	987	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCT	1046
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	1047	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG	1106
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1107	ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC	1166
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1167	TTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT	1226
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACCGGCATGAACCGGCCCTTCCTGC	785
Db	1227	TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGC	1286
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1287	TTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA-	1345
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1346	-----GCCCTGGACACCAACTATTGCT	1367
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1368	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC	1427
Qy	966	TGGGCTGGAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1428	TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1487
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1488	GCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC	1547
Qy	1086	ACAACCCGGGCGCGTCCGGCGGCGCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA	1145
Db	1548	ATAACCCGGGCGCCTCGGCGGCGCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCA	1607
Qy	1146	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT	1205
Db	1608	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT	1667

[illegible]

RESULT 2

US-10-037-270-220

; Sequence 220, Application US/10037270

: Publication No. US20030104529A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunging

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

APPLICANT: Tillinghast, John

APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/10/037,270

; CURRENT FILING DATE: 2002-01-04

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

```

; NUMBER OF SEQ ID NOS: 1104

```

```

; SOFTWARE: pt FL genes Version 1.0

```

; SEO ID NO 220

; LENGTH: 2742

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

LOCATION: (842) .. (2014)

US-10-037-270-220

Query Match 72.8%; Score 981.4; DB 14; Length 2742;

Best Local Similarity 85.5%; Pred. No. 5.1e-254;

Matches 1147; Conservative 0; Mismatches 146; Indels 48; Gaps 3;

```
Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy     66 TAGTGCTGACGCCTGGCCGGCCGGCCGGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    897 TGGTGCTGACGCCTGGCCGGCCGGCCGGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy    126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy    186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG 245
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCG 1076

Qy    246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy    306 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC 365
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1137 CTGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAACCCACAACG 1196

Qy    366 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy    426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1257 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGCAGAGCTGCGTCTGCTGA 1316

Qy    486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1317 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCT 1376

Qy    546 GGCCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG 605
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1377 GGCATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTG 1436

Qy    606 ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCC 665
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1437 ATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCC 1496

Qy    666 TCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT 725
      | || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1497 TTAGCGCCCACTGCTCTCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGT 1556

Qy    726 TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCTCTGC 785
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1557 TCACTACCGGCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCTTCTCTGC 1616

Qy    786 TCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG 845
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   1617 TTCTCATGGCCACCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGA- 1675
```

```

Qy      846 CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT 905
          ||||| ||||| |||
Db      1676 -----GCCCTGGACACCAACTATTGCT 1697

Qy      906 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC 965
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1698 TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACC 1757

Qy      966 TGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT 1025
          | ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1758 TCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGCCTCGGGCCCT 1817

Qy      1026 GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC 1085
          | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1818 GCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGC 1877

Qy      1086 ACAACCCGGGCGCGTCGGCGGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCA 1145
          | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1878 ATAACCCGGGCGCCTCGGCGGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCA 1937

Qy      1146 TCGTGTAACGCTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTT 1205
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1938 TCGTGTAACGCTGGGCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCT 1997

Qy      1206 CCTGCAAGTGCAGCTGAGGCCCGCCCGCCCAAGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 1265
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1998 CCTGCAAGTGCAGCTGAGGTCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 2057

Qy      1266 CCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACA-TCGTGCCCCAA 1316
          ||| |||| ||| || || ||||| ||||| ||||| ||||| ||||| |||||
Db      2058 CCCGCCCCGCCCCGCTGCCTTGCCCATGGGGGCTGTATTTAAGGACACCCGTGCCCCAA 2117

Qy      1317 GCCCACTTGGGATCGATTAAA 1337
          ||||| |||| | | ||
Db      2118 GCCCACCTGGGGCCCCATTAA 2138

```

RESULT 3

US-10-087-268-4

; Sequence 4, Application US/10087268

; Publication No. US20030119010A1

; GENERAL INFORMATION:

; APPLICANT: Jonsonn, Julie Ruth

; APPLICANT: Powell, Elizabeth Ellen

; TITLE OF INVENTION: Polypeptides and polynucleotides linked to a disease or condition

; FILE REFERENCE: Fibrosis

; CURRENT APPLICATION NUMBER: US/10/087,268

; CURRENT FILING DATE: 2002-03-01

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1821

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

```
; NAME/KEY: 5'UTR
; LOCATION: (1)..(511)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (512)..(1684)
; OTHER INFORMATION:
; NAME/KEY: sig_peptide
; LOCATION: (512)..(598)
; OTHER INFORMATION:
; NAME/KEY: 3'UTR
; LOCATION: (1685)..(1821)
; OTHER INFORMATION:
```

Query Match 72.7%; Score 980.2; DB 14; Length 1821;
Best Local Similarity 85.6%; Pred. No. 9.8e-254;
Matches 1133; Conservative 0; Mismatches 143; Indels 47; Gaps 2;

Qy	6	CCGAGATGGCGCCTTCGGGGCTGCGGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC	65
Db	507	CCCCCATGCGGCCCTCCGGGCTGCGGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC	566
Qy	66	TAGTGCTGACGCCTGGCCGGCCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	567	TGGTGCTGACGCCTGGCCCGCCGGCCGCGGGACTATCCACCTGCAAGACTATCGACATGG	626
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	627	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC	686
Qy	186	TTGCCAGCCCCCGAGCCAGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG	245
Db	687	TCGCCAGCCCCCGAGCCAGGGGAGGTGCCGCCCGGCCGCTGCCCAGGCCGTGCTCG	746
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGAAACCGAGCCCGAGC	305
Db	747	CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAAAACCGAGCCCGAGC	806
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAAGCGGCAACC	365
Db	807	CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGAAACCCACAACG	866
Qy	366	AAATCTATGATAAATTC AAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	867	AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG	926
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	927	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	986
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT	545
Db	987	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT	1046
Qy	546	GGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGCGCTGTCCTTTG	605
Db	1047	GGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGGTTATCTTTTG	1106


```
; Publication No. US20030050265A1
; GENERAL INFORMATION:
; APPLICANT: Nicholas M. Dean
; APPLICANT: Susan F. Murray
; TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH
; TITLE OF INVENTION: FACTOR BETA EXPRESSION
; FILE REFERENCE: ISPH-0607
; CURRENT APPLICATION NUMBER: US/09/948,002
; CURRENT FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 09/661,753
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/154,546
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 71
; SEQ ID NO 28
; LENGTH: 2745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (842)...(2017)
US-09-948-002-28
```

```
Query Match          72.5%; Score 977.8; DB 11; Length 2745;
Best Local Similarity 85.3%; Pred. No. 4.8e-253;
Matches 1146; Conservative 0; Mismatches 147; Indels 50; Gaps 3;
```

```
Qy      6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65
      || ||| |||| || ||||||||||| ||||||||||| ||||||||||| ||
Db      837 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTAC 896

Qy      66 TAGTGCTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG 125
      | ||||||||||||||| ||||||||| ||||| ||||||||||||||| |||||||||
Db      897 TGGTGCTGACGCCTGGCCCCCGGCCGCGGGACTATCCACCTGCAAGACTATCGACATGG 956

Qy      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      ||||||||||||||| ||||||||||| ||||||||||| ||||||||||| ||||
Db      957 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 1016

Qy      186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      | ||||||||||||||| ||||||||||| ||||||||||| ||||||| || |||
Db      1017 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 1076

Qy      246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGAAAGTGTCGAACCGGAGCCCGAGC 305
      | || ||||||||| ||||||||||||| ||||||||| ||||| |||||||||||||
Db      1077 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 1136

Qy      306 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 365
      | ||||| ||||||||||||||||||| ||||||||||| ||||||||| || ||||
Db      1137 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG 1196

Qy      366 AAATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      ||||||||| || ||||||| | ||||||| ||||||| ||||||||| || |
Db      1197 AAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 1256

Qy      426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT-- 483
      ||||||| ||||||||| || ||||| || ||||||||| ||||||||||||| |||||
```


Db	1257	AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA	1316
Qy	484	-GAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATT	542
Db	1317	GGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATT	1376
Qy	543	CCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTACCGGAGTGGCTGTCTT	602
Db	1377	CCTGGCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTT	1436
Qy	603	TTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTC	662
Db	1437	TTGATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTC	1496
Qy	663	GCCTCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACG	722
Db	1497	GCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACG	1556
Qy	723	GGTTC AATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCC	782
Db	1557	GGTTC ACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCC	1616
Qy	783	TGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCC	842
Db	1617	TGCTTCTCATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCC	1676
Qy	843	GAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACT	902
Db	1677	GA-----GCCCTGGACACCAACTATT	1697
Qy	903	GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGG	962
Db	1698	GCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGG	1757
Qy	963	ACCTGGGCTGGAAGTGGATT CATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGC	1022
Db	1758	ACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGC	1817
Qy	1023	CCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACC	1082
Db	1818	CCTGCCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACC	1877
Qy	1083	AGCACAACCCGGGCGCGTGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGC	1142
Db	1878	AGCATAACCCGGGCGCCTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGC	1937
Qy	1143	CCATCGTG TACTACGTGGGCGCGAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1202
Db	1938	CCATCGTG TACTACGTGGGCGCGAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC	1997
Qy	1203	GTCCTGCAAGTGCAGCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCC	1262
Db	1998	GTCCTGCAAGTGCAGCTGAGGTCCCGCCCCGCCCGCCCCGCCCGGCGAGCCCCGGCCC	2057
Qy	1263	GGCCCCACCCCGCCCGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCC	1314
Db	2058	CACCCCGCCCCGCCCGCCTTGCCCATGGGGGCTGTATTTAAGGACACCGTGCCCC	2117

Qy 1315 AAGCCCACTTGGGATCGATTAAA 1337
 ||||| ||| | ||
 Db 2118 AAGCCCACTTGGGGCCCCATTAA 2140

RESULT 5

US-09-948-002-1

; Sequence 1, Application US/09948002

; Publication No. US20030050265A1

; GENERAL INFORMATION:

APPLICANT: Nicholas M. Dean

; APPLICANT: Susan F. Murray

TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH

; TITLE OF INVENTION: FACTOR BETA EXPRESSION

FILE REFERENCE: ISPH-0607

; CURRENT APPLICATION NUMBER: US/09/948,002

; CURRENT FILING DATE: 2000-09-05

; PRIOR APPLICATION NUMBER: 09/661,753

; PRIOR FILING DATE: 2000-09-14

PRIOR APPLICATION NUMBER: 60/154,546

PRIOR FILING DATE: 1999-09-17

; NUMBER OF SEQ ID NOS: 71

; SEO ID NO 1

; LENGTH: 2094

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

LOCATION: (868) ... (2040)

US-09-948-002-1

Query Match 63.7%; Score 859; DB 11; Length 2094;
Best Local Similarity 81.5%; Pred. No. 4.1e-221;
Matches 1038; Conservative 0; Mismatches 190; Indels 45; Gaps 2;

Qy	6	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGTGCTGCCGCTGCTGTGGCTGC	65
Db	863	CCCCCATGCCGCCCTCGGGGCTGCGGCTACTGCCGCTTCTGCTCCCACTCCCGTGGCTTC	922
Qy	66	TAGTGCTGACGCCTGGCCGGCCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	923	TAGTGCTGACGCCCCGGGAGGCCAGCCGCGGGACTCTCCACCTGCAAGACCATCGACATGG	982
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	983	AGCTGGTGAAACGGAAGCGCATCGAAGCCATCCGTGGCCAGATCCTGTCCAAACTAAGGC	1042
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	1043	TCGCCAGTCCCCCAAGCCAGGGGGAGGTACCGCCCGGCCCGCTGCCCGAGGCGGTGCTCG	1102
Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC	305
Db	1103	CTTTGTACAACAGCACCCGCGACCGGGTGGCAGGCGAGAGCGCCGACCCAGAGCCGGAGC	1162
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACC	365

Db	1163	CCGAAGCGGACTACTATGCTAAAGAGGTCAACCGCGTGCTAATGGTGGACCGCAACAACG	1222
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	1223	CCATCTATGAGAAAAACAAAGACATCTCACACAGTATATATATGTTCTTCAATACGTCAG	1282
Qy	426	AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	1283	ACATTCGGGAAGCAGTGCCCGAACCCCCATTGCTGTCCCGTGCAGAGCTGCGCTTGCAGA	1342
Qy	486	GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCCT	545
Db	1343	GATTAAAAATCAAGTGTGGAGCAACATGTGGAACCTCTACCAGAAATATAGCAACAATTCCCT	1402
Qy	546	GGCGTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCTTTTG	605
Db	1403	GGCGTTACCTTGGTAAACCGGCTGCTGACCCCCACTGATACGCCTGAGTGGCTGTCTTTTG	1462
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1463	ACGTCACCTGGAGTTGTACGGCAGTGGCTGAACCAAGGAGACGGAATACAGGGCTTTCGAT	1522
Qy	666	TCAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1523	TCAGCGCTCACTGCTCTTGTGACAGCAAAGATAACAAACTCCACGTGGAAATCAACGGGA	1582
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGC	785
Db	1583	TCAGCCCCAAACGTCGGGGCGACCTGGGCACCATCCATGACATGAACCGGCCCTTCCTGC	1642
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1643	TCCTCATGGCCACCCCCCTGGAAAGGGCCAGCACCTGCACAGCTCACGGCACCGGAGA-	1701
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1702	-----GCCCTGGATACCAACTATTGCT	1723
Qy	906	TCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1724	TCAGCTCCACAGAGAAGAACTGCTGTGTGCGGCAGCTGTACATTGACTTTAGGAAGGACC	1783
Qy	966	TGGGCTGGAAGTGGATTGATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1784	TGGGTTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTCTGTCTGGGACCCT	1843
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1844	GCCCCATATTTGGAGCCTGGACACACAGTACAGCAAGGTCCTTGCCCTCTACAACCAAC	1903
Qy	1086	ACAACCCGGGCGCGTCCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA	1145
Db	1904	ACAACCCGGGCGCTTCGGCGTCAACGTGCTGCGTGCCGCAGGCTTTGGAGCCACTGCCCA	1963
Qy	1146	TCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGC GTT	1205

```
Db      1964 TCGTCTACTACGTGGGTCGCAAGCCCAAGGTGGAGCAGTTGTCCAACATGATTGTGCGCT 2023
Qy      1206 CCTGCAAGTG CAGCTGAGGCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGC 1265
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2024 CCTGCAAGTG CAGCTGAAGCCCCGCCCGC-----CCCGCCCCTCCCGGCAGGCCCGGC 2077
Qy      1266 CCCACCCCCGCC 1278
        ||| ||||| ||||| |||||
Db      2078 CCCGCCCCCCGCC 2090
```

RESULT 6

US-09-948-002-27

; Sequence 27, Application US/09948002

: Publication No. US20030050265A1

: GENERAL INFORMATION:

APPLICANT: Nicholas M. Dean

; APPLICANT: Susan F. Murray

: TITLE OF INVENTION: ANTISENSE MODULATION OF TRANSFORMING GROWTH

: TITLE OF INVENTION: FACTOR BETA EXPRESSION

FILE REFERENCE: ISPH-0607

; CURRENT APPLICATION NUMBER: US/09/948,002

; CURRENT FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: 09/661,753

; PRIOR FILING DATE: 2000-09-14

PRIOR APPLICATION NUMBER: 60/154,546

; PRIOR FILING DATE: 1999-09-17

NUMBER OF SEQ ID NOS: 71

; SEQ ID NO 27

; LENGTH: 1585

; TYPE: DNA

; ORGANISM: Rattus norvegicus

FEATURE:

; NAME/KEY: CDS

LOCATION: (413)...(1585)

US-09-948-002-27

Query Match 62.2%; Score 839.4; DB 11; Length 1585;

Best Local Similarity 81.9%; Pred. No. 7.2e-216;

Matches 997; Conservative 0; Mismatches 181; Indels 39; Gaps 1;

Qy	6	CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC	65
Db	408	CCCCCATGCCGCCCTCGGGGCTGCGGCTGCTGCCGCTTCTGCTCCCACTCCCGTGGCTTC	467
Qy	66	TAGTGCTGACGCCTGGCCGGCCGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG	125
Db	468	TAGTGCTGACGCCCGGGAGGCCAGCCGCGGACTCTCCACCTGCAAGACCATCGACATGG	527
Qy	126	AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC	185
Db	528	AGCTGGTGAACCGGAAGCGCATCGAAGCCATCCGTGGCCAGATCCTGTCCAAACTAAGGC	587
Qy	186	TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGG	245
Db	588	TCGCCAGTCCCCCGAGCCAGGGGGAGGTACCGCCGGGGCCCGCTGCCCGAGGCGGTGCTCG	647

Qy	246	CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTCGAACCGGAGCCCGAGC	305
Db	648	CTTTGTACAACAGCACCCGCGACCGGGTGGCAGGCGAGAGCGCTGACCCGGAGCCCGAGC	707
Qy	306	CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAAGCGGCAACC	365
Db	708	CCGAGGCGGACTACTACGCCAAAGAAGTCACCCGCGTGCTAATGGTGGACCGCAACAACG	767
Qy	366	AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG	425
Db	768	CAATCTATGACAAAACCAAAGACATCACACACAGTATATATATGTTCTTCAATACGTCAG	827
Qy	426	AGCTCCGGGAAGCGGTGCCGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA	485
Db	828	ACATTCGGGAAGCAGTGCCAGAACCCCCATTGCTGTCCCGTGCAGAGCTGCGCCTGCAGA	887
Qy	486	GGCTCAAGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCT	545
Db	888	GATTCAAGTCAACTGTGGAGCAACACGTAGAACTCTACCAGAAATATAGCAACAATTCCT	947
Qy	546	GGCGTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTG	605
Db	948	GGCGTTACCTTGGTAAACCGGCTGCTGACCCCCACTGATACGCCTGAGTGGCTGTCTTTTG	1007
Qy	606	ATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCC	665
Db	1008	ACGTCACTGGAGTTGTCCGGCAGTGGCTGAACCAAGGAGACGGAATACAGGGCTTTTCGCT	1067
Qy	666	TCAGTGCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGT	725
Db	1068	TCAGTGCTCACTGCTCTTGTGACAGCAAAGATAATGTACTCCACGTGGAAATCAATGGGA	1127
Qy	726	TCAATTCTGGCCGCCGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCTCTGC	785
Db	1128	TCAGTCCCAAACGTGAGGTGACCTGGGCACCATCCATGACATGAACCGACCCTTCTCTGC	1187
Qy	786	TCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAG	845
Db	1188	TCCTCATGGCCACCCCCCTGGAAAGGGCTCAACACCTGCACAGCTCCAGGCACCGGAGA-	1246
Qy	846	CCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCT	905
Db	1247	-----GCCCTGGATACCAACTACTGCT	1268
Qy	906	TCAGTCCACGGAGAAGAAGTCTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACC	965
Db	1269	TCAGTCCACAGAGAAGAAGTCTGTGTACGGCAGCTGTACATTGACTTTAGGAAGGACC	1328
Qy	966	TGGGCTGGAAGTGGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCT	1025
Db	1329	TGGGTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGTCTGGGGCCCT	1388
Qy	1026	GTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGC	1085
Db	1389	GCCCCACATTTGGAGCCTGGACACACAGTACAGCAAGGTCCTTGCCCTCTACAACCAAC	1448
Qy	1086	ACAACCCGGGCGCGTCCGGCGCGCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCA	1145

Qy	11	ATGGCGCCTTCGGGGCTGCGGCTCTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTG	70
Db	1	ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTG	60
Qy	71	CTGACGCCTGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTG	130
Db	61	CTGACGCCTGGCCGGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGGAGCTG	120
Qy	131	GTGAAGCGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCC	190
Db	121	GTGAAGCGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCC	180
Qy	191	AGCCCCCAGCCAGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGGCTCTT	250
Db	181	AGCCCCCAGCCAGGGGAGGTGCCGCCCGGCCGCTGCCGAGGCCGTGCTCGCCCTG	240
Qy	251	TACAACAGTACCCGCGACCGGGTAGCCGGGAAAGTGTCGAACCGGAGCCCGAGCCAGAG	310

Db	241	TACAACAGCACCCGCGACCCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAG	300
Qy	311	GCGGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAAGCGGCAACCAAATC	370
Db	301	GCCGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAACCAACAACGAAATC	360
Qy	371	TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGGCTC	430
Db	361	TATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGGCTC	420
Qy	431	CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT--GAGG	487
Db	421	CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGAGG	480
Qy	488	CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGG	547
Db	481	CTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGG	540
Qy	548	CGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGTCTTTTGAT	607
Db	541	CGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGAT	600
Qy	608	GTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTC	667
Db	601	GTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTT	660
Qy	668	AGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAAATTAACGGGTTC	727
Db	661	AGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTC	720
Qy	728	AATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTC	787
Db	721	ACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTT	780
Qy	788	CTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGC	829
Db	781	CTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGC	822

US-09-756-283A-21

; Sequence 21, Application US/09756283A

; GENERAL INFORMATION:

APPLICANT: Dreja, Hanna Stina

; TITLE OF INVENTION: Latent Fusion Protein

; CURRENT APPLICATION NUMBER: US/09/756,283A

; NUMBER OF SEQ ID NOS: 100

; SEQ ID NO 21

TYPE: DNA

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: mIFNbeta-LAP construct
; NAME/KEY: CDS
; LOCATION: (1)..(1344)
US-09-756-283A-21
```

```
Query Match          43.8%;  Score 591.4;  DB 10;  Length 1352;
Best Local Similarity 87.0%;  Pred. No. 3.7e-149;
Matches 662;  Conservative 0;  Mismatches 96;  Indels 3;  Gaps 1;
```

```
Qy      73  GACGCCTGGCCGGCCGGCCGGCCGGACTGTCCACCTGCAAGACCATCGACATGGAGCTGGT 132
      ||| ||| ||| | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      582  GCGCGGAGGGGGCTCAGCGGCCGCACTATCCACCTGCAAGACTATCGACATGGAGCTGGT 641

Qy      133  GAAGCGGAAGCGCATCGAGGCCATTGCGGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAG 192
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      642  GAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAG 701

Qy      193  CCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTA 252
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      702  CCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTGTA 761

Qy      253  CAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAGGC 312
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      762  CAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGC 821

Qy      313  GGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAAGCGGCAACCAAATCTA 372
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      822  CGACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAACCCACAACGAAATCTA 881

Qy      373  TGATAAATTCAAGGGCACCCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCG 432
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      882  TGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCG 941

Qy      433  GGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCT--GAGGCT 489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      942  AGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGAGGCT 1001

Qy      490  CAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCG 549
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1002  CAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCG 1061

Qy      550  CTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCCTTTGATGT 609
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1062  ATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTATGT 1121

Qy      610  CACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTCGCCTCAG 669
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1122  CACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAG 1181

Qy      670  TGCCCCACTCTTCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAA 729
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1182  CGCCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTTCA 1241

Qy      730  TTCTGGCCGCCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCCTGCTCCT 789
      | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

Db      1242 TACCGGCCGCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCTTTCTGCTTCT 1301
Qy      790 CATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCT 830
        |||||||||||||||||||||||||||||||
Db      1302 CATGGCCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCT 1342

```

RESULT 9

```

US-09-911-904-167
; Sequence 167, Application US/09911904
; Publication No. US20030096234A1
; GENERAL INFORMATION:
; APPLICANT: Farr, Spencer B.
; APPLICANT: Pickett, Gavin G.
; APPLICANT: Neft, Robin Eileen
; APPLICANT: Dunn, II, Robert Thomas
; TITLE OF INVENTION: CANINE TOXICITY GENES
; FILE REFERENCE: 400742000200
; CURRENT APPLICATION NUMBER: US/09/911,904
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/220,057
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 167
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(489)
; OTHER INFORMATION: n = A, T, C or G
US-09-911-904-167

```

```

Query Match          27.2%; Score 366.4; DB 11; Length 489;
Best Local Similarity 86.1%; Pred. No. 1e-88;
Matches 445; Conservative 0; Mismatches 32; Indels 40; Gaps 2;

```

```

Qy      774 GGCCCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCC 833
        | |||||||||||||||||||||||||||||||||||||||||||
Db      1   GACCCTTCCTGCTCCTCATGGCCACCCCACTGGAGAGGGCCCAGCACCTGCACAGCTCCC 60

Qy      834 GGCACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATA 893
        |||| |||||                                     |||||||||
Db      61  GGCAGCGCCG-----GGCCCTGGACA 81

Qy      894 CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT 953
        |||||||||||||||||||||||||||||||||||||||||||
Db      82  CCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACT 141

Qy      954 TCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCT 1013
        |||| ||||| |||||||||||||||||| |||| ||||| || || |||
Db      142 TCCGCAAGGATCTGGGCTGGAAGTGGATCCATGAGCCCAAGGGTTACCACGCTAATTCT 201

Qy      1014 GCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTC 1073
        |||||||||||||| ||||||| ||||||| ||||| |||||||||||||||
Db      202 GCCTGGGGCCCTGCCCCTACATTTGGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCC 261

```

Qy 1074 TGTACAACCAGCACAAACCGGGCGCGTTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 1133
 |||
 Db 262 TGTACAACCAGCACAAACCGGGCGCGTTCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGG 321
 Qy 1134 AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACA 1193
 |||
 Db 322 AGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCAACA 381
 Qy 1194 TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCGCCCCCG-CCCACAGCCCCGCCACC 1252
 |||
 Db 382 TGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCGCCCCCGTCCGGCAGGCCCGCCCCAC 441
 Qy 1253 CGGCAGGCCCGGCCCAACCCCGCCCGCTACCGGG 1289
 |||
 Db 442 CGGCAGGNCCGGCCCCCGCCCGCTGCGCCGGG 478

RESULT 10

US-09-813-271B-1

; Sequence 1, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

; APPLICANT:

; (A) Nico Cerletti

; TITLE OF INVENTION: New process for the production of
 ; biologically active protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. US20020115834A1artis Patent Department

; STREET: 564 Morris Avenue

; CITY: Summit

; STATE: New Jersey

; COUNTRY: USA

; ZIP: 07901

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/813,271B

; FILING DATE: 20-Mar-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP95/02719

; FILING DATE: 12-Jul-95

; APPLICATION NUMBER: EPO 94810439.3

; FILING DATE: 25-Jul-94

; ATTORNEY/AGENT INFORMATION:

; NAME: Pfeiffer, Hesna J. .

; REGISTRATION NUMBER: 22640

; REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (908) 522-6940

; TELEFAX: (908) 522-6955

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; SEQ ID NO 3
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (254) ... (1492)
US-09-906-158-3

Query Match 17.7%; Score 238.4; DB 11; Length 2574;
Best Local Similarity 53.6%; Pred. No. 4.1e-54;
Matches 668; Conservative 0; Mismatches 521; Indels 57; Gaps 6;

```
Qy      36  TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 95
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     261  TGCACTTGCAAAGGGCTCTGGTGGTCTGGCCCTGCTGAACTTTGCCACGGTCAGCCTCT 320

Qy      96  GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 155
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     321  CTCTGTCCACTTGCAACACCTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGGAAGCCA 380

Qy     156  TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 215
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     381  TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCAACGGTGATGA 440

Qy     216  CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 275
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     441  CCCACG-----TCCCCTATCAGGTCTTGGCCCTTTACAACAGCACCCGGGAGCTGCTGG 494

Qy     276  CCGGGGAAAAGTGTGCAACCGGAGCCCG-----AGCCAGAGGCGGACTACT 320
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     495  AGGAGATGCATGGGGAGAGGGAGGAAGGCTGCACCCAGGAAAACACCGAGTCGGAATACT 554

Qy     321  ACGCCAAGGAGGTCACCCGGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT 380
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     555  ATGCCAAAGAAATCCATAAATTCGACATGATCCAGGGGCTGGCGGAGCACAACGAACCTGG 614

Qy     381  TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGG 440
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     615  CTGTCTGCCCTAAAGGAATTACCTCCAAGTTTTCCGCTTCAATGTGTCTCAGTGGAGA 674

Qy     441  TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAG 500
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     675  AAAATAGAACCAACCTATTCCGAGCAGAATTCCGGGTCTTGCGGGTGCCCAACCCAGCT 734

Qy     501  TGGAGCAGCACGTGGAG-----CTATACCAGAAATACAGCAATG 539
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     735  CTAAGCGGAATGAGCAGAGGATCGAGCTCTTCAGATCCTTCGGCCAGATGAGCACATTG 794

Qy     540  ATTCCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGT 599
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     795  CCAAACAGCGCTATATCGGTGGCAAGAATCTGCCACACGGGGCACTGCCGAGTGGCTGT 854

Qy     600  CCTTTGATGTCAACCGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTT 659
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     855  CCTTTGATGTCACTGACACTGTGCGTGAGTGGCTGTTGAGAAGAGAGTCCAACTTAGGTC 914
```

QY 660 TTCGCCTCAGTGCCCACTCTTCTCTGACAGCAAAGATAACA-----CACTCCACGTGG 713
 | | | | | | | | | | | | | | | | | | | | | |
 Db 915 TAGAAATCAGCATTCACTGTCCATGTACACCTTTAGCCCAATGGAGATATCCTGGAAA 974

QY 714 AAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACC 773
 | | | | | | | | | | | | | | | | | | | | | |
 Db 975 ACATTCACGAGGTGATGGAAATCAAATTCAAAGGCGTGGACAATGAGGATGACCATGGCC 1034

QY 774 GGCCCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCC 833
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1035 GTGGAGATCTGGGGCGCCTCAAGAAGCAGAAGGATCACCACAACCCTCATCTAATCCTCA 1094

QY 834 GG-----CACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAA--GG 884
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1095 TGATGATTCCCCCACACCGGCTCGACAACCCGGGCCAGGGGGGTGAGAGGAAGAAGCGGG 1154

QY 885 CCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCT 944
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1155 CTTTGGACACCAATTACTGCTTCCGCAACTTGGAGGAGAACTGCTGTGTGCGCCCCCTCT 1214

QY 945 ACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATG 1004
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1215 ACATTGACTTCCGACAGGATCTGGGCTGGAAGTGGGTCCATGAACCTAAGGGCTACTATG 1274

QY 1005 CCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGG 1064
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1275 CCAACTTCTGCTCAGGCCCTTGCCCATACCTCCGCAGTGCAGACACAACCCACAGCACGG 1334

QY 1065 TCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGC 1124
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1335 TGCTGGGACTGTACAACACTCTGAACCCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCCC 1394

QY 1125 AGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGCGCAAGCCCAAGGTGGAGCAGC 1184
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1395 AGGACCTGGAGCCCCTGACCATCCTGTACTATGTTGGGAGGACCCCAAGTGGAGCAGC 1454

QY 1185 TGTCCAACATGATCGTGCCTTCTGCAAGTGCAGCTGAGGCCCCGC 1230
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1455 TCTCCAACATGGTGGTGAAGTCTTGTAATGTAGCTGAGACCCAC 1500

RESULT 12

US-10-028-158-20

; Sequence 20, Application US/10028158

; Publication No. US20020110833A1

; GENERAL INFORMATION:

; APPLICANT: Caniggia, Isabella

; APPLICANT: Post, Martin

; APPLICANT: Lye, Stephen

; TITLE OF INVENTION: METHODS TO DIAGNOSE A REQUIRED REGULATION OF

; TITLE OF INVENTION: TROPHOBLAST

; FILE REFERENCE: 11757.38USWO

; CURRENT APPLICATION NUMBER: US/10/028,158

; CURRENT FILING DATE: 2001-12-20

; PRIOR APPLICATION NUMBER: US/09/380,662

; PRIOR FILING DATE: 1999-12-21

; PRIOR APPLICATION NUMBER: PCT/CA98/00180

; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: US 60/039,919
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 2574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (254)..(1492)
US-10-028-158-20

Query Match 17.7%; Score 238.4; DB 13; Length 2574;
Best Local Similarity 53.6%; Pred. No. 4.1e-54;
Matches 668; Conservative 0; Mismatches 521; Indels 57; Gaps 6;

```
Qy      36 TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 95
      ||| ||| || ||| || ||| ||||| || | | | |
Db     261 TGCACTTGCAAAGGGCTCTGGTGGTCCTGGCCCTGCTGAACCTTGCCACGGTCAGCCTCT 320

Qy      96 GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 155
      ||||| ||| ||| ||| ||| || | ||| ||| | | |||
Db     321 CTCTGTCCACTTGCAACACCTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGAAGCCA 380

Qy     156 TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 215
      || | || ||||| || ||||| ||| ||||| ||| | || |
Db     381 TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCAACGGTGATGA 440

Qy     216 CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 275
      | | || | | | | | | | | | | | | | | | | |
Db     441 CCCACG-----TCCCTATCAGGTCTGGCCCTTTACAACAGCACCCGGGAGCTGCTGG 494

Qy     276 CCGGGGAAAGTGTCGAACCGGAGCCCG-----AGCCAGAGGCGGACTACT 320
      | | | | | | | | | | | | | | | | | | | |
Db     495 AGGAGATGCATGGGGAGAGGGAGGAAGGCTGCACCCAGGAAAACACCGAGTCGGAATACT 554

Qy     321 ACGCCAAGGAGGTCACCCGCTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT 380
      | ||||| || | | | | | | | | | | | | | |
Db     555 ATGCCAAAGAAATCCATAAATTCCGACATGATCCAGGGGCTGGCGGAGCACAAACGAATGG 614

Qy     381 TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGG 440
      || | | | | | | | | | | | | | | | | | |
Db     615 CTGTCTGCCCTAAAGGAATTACCTCCAAGGTTTTCCGCTTCAATGTGTCCTCAGTGGAGA 674

Qy     441 TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAG 500
      | | | | | | | | | | | | | | | | | | |
Db     675 AAAATAGAACCAACCTATTCCGAGCAGAATTCCGGGTCTTGCGGTGCCCAACCCAGCT 734

Qy     501 TGGAGCAGCACGTGGAG-----CTATACCAGAAATACAGCAATG 539
      ||| | | | | | | | | | | | | | | | |
Db     735 CTAAGCGGAATGAGCAGAGGATCGAGCTCTTCCAGATCCTTCGGCCAGATGAGCACATTG 794

Qy     540 ATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGA CTACCGGAGTGGCTGT 599
      ||||| || | | | | | | | | | | | | | |
Db     795 CCAAACAGCGCTATATCGGTGGCAAGAATCTGCCACACGGGGCACTGCCGAGTGGCTGT 854
```

Qy 600 CCTTTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTT 659
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 855 CCTTTGATGTCAGTACTGACACTGTGCGTGAGTGGCTGTTGAGAAGAGAGTCCAACCTTAGGTC 914
 Qy 660 TTCGCCTCAGTGCCCACTCTTCTCTGACAGCAAAGATAACA-----CACTCCACGTGG 713
 | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 915 TAGAAATCAGCATTCACTGTCCATGTACACCTTTGAGCCCAATGGAGATATCCTGGAAA 974
 Qy 714 AAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACC 773
 | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 975 ACATTCACGAGGTGATGGAAATCAAATTCAAAGGCGTGGACAATGAGGATGACCATGGCC 1034
 Qy 774 GGCCCTTCCTGCTCCTCATGGCCACCCGCTGGAGAGGGCCAGCACCTGCACAGCTCCC 833
 | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 1035 GTGGAGATCTGGGGCGCCTCAAGAAGCAGAAGGATCACCACAACCCTCATCTAATCCTCA 1094
 Qy 834 GG-----CACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGACGACAA---GG 884
 | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 1095 TGATGATTCCCCACACCGGCTCGACAACCCGGGCCAGGGGGGTGAGAGGAAGAAGCGGG 1154
 Qy 885 CCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCT 944
 | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 1155 CTTTGGACACCAATTACTGCTTCGCAACTTGGAGGAGAACTGCTGTGTGCGCCCCCTCT 1214
 Qy 945 ACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTGATGAACCCAAGGGCTACCATG 1004
 | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 1215 ACATTGACTTCCGACAGGATCTGGGCTGGAAGTGGGTCCATGAACCTAAGGGCTACTATG 1274
 Qy 1005 CCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGG 1064
 | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 1275 CCAACTTCTGCTCAGGCCCTTGCCCATACCTCCGCAGTGCAGACACAACCCACAGCACGG 1334
 Qy 1065 TCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGC 1124
 | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 1335 TGCTGGGACTGTACAACACTCTGAACCCTGAAGCATCTGCCTCGCCTTGCTGCGTGCCCC 1394
 Qy 1125 AGGCGCTGGAGCCACTGCCCATCGTGTAACGCTGGGCGCGCAAGCCCAAGGTGGAGCAGC 1184
 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1395 AGGACCTGGAGCCCCTGACCATCCTGTACTATGTTGGGAGGACCCCAAGTGGAGCAGC 1454
 Qy 1185 TGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCCCCGC 1230
 | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1455 TCTCCAACATGGTGGTGAAGTCTTGTAATGTAGCTGAGACCCAC 1500

RESULT 13

US-09-957-458B-9

; Sequence 9, Application US/09957458B

; Publication No. US20030166271A1

; GENERAL INFORMATION:

; APPLICANT: Chen, Una

; TITLE OF INVENTION: Method for growing stem cells

; FILE REFERENCE: P66567US0

; CURRENT APPLICATION NUMBER: US/09/957,458B

; CURRENT FILING DATE: 2001-09-21

; PRIOR APPLICATION NUMBER: PCT/EP00/08247

; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: EP 99116533
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 4382
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector for transforming supporting cell with a foreign to
express
; OTHER INFORMATION: a gene product of interest
US-09-957-458B-9

Query Match 17.3%; Score 233.2; DB 12; Length 4382;
Best Local Similarity 52.8%; Pred. No. 1.2e-52;
Matches 660; Conservative 0; Mismatches 533; Indels 57; Gaps 5;

```
Qy      26 CTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGG 85
      |||| || | ||| ||| || |||| | ||| ||||| |||
Db      455 CTGCAGCCCATGCACTTGCAAAGGGCTCTGGTAGTCCTGGCCCTGCTGAACTTGGCCACA 514

Qy      86 CCGGCCGCCGACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGC 145
      | | | ||||| |||| | | | ||| | | | ||| |||
Db      515 ATCAGCCTCTCTCTGTCCACTTGCAACACGTTGGACTTCGGCCACATCAAGAAGAAGAGG 574

Qy     146 ATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAG 205
      | || ||||| | || |||| | || |||| | || | ||||| |||
Db      575 GTGGAAGCCATTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCA 634

Qy     206 GGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGC 265
      || | | | || | || | | || |||| | ||||| |||||
Db      635 TCGGTGATGACCCACG-----TCCCCTATCAGGTCTTGGCACTTTACAACAGCACCCGG 688

Qy     266 GACCGGGTAGCCG-----GGGAAAGTGTCGAACCGGAGCCCCGAGCCAGAG 310
      || | | | | |||| | | | | | || | |||
Db      689 GAGTTGCTGGAAGAGATGCACGGGGAGAGGGAGGAAGGCTGCACTCAGGAGACCTCGGAG 748

Qy     311 GCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATC 370
      | || |||| | |||| | || | || | | | | | |||
Db      749 TCTGAGTACTATGCCAAAGAGATCCATAAATTGACATGATCCAGGGACTGGCGGAGCAC 808

Qy     371 TATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTC 430
      |||| | || || | || | || | || | || | || | |||
Db      809 AATGAACTGGCCGTCTGCCCCAAAGGAATTACCTCTAAGGTTTTTCGTTTCAATGTGTCC 868

Qy     431 CGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTC 490
      | | | | | || | || || | | | | | || | |||
Db      869 TCAGTGAGAAAAATGGAACCAATCTGTTCCGGGCAGAGTTCGGGTCTTGCGGTGCC 928

Qy     491 A-----AGTTAAAAGTGAGCAGCACGTGGAGCTATACCAGAAA 529
      | || | | |||| | | | |||
Db      929 AACCCAGCTCCAAGCGCACAGAGCAGAGAATTGAGCTCTTCCAGATACTTCGACCGGAT 988

Qy     530 TACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCAGTGACTCACCG 589
      | || | ||||| | | | | ||||| |||
```



```

; CURRENT APPLICATION NUMBER: US/09/906,158
; CURRENT FILING DATE: 2001-07-14
; NUMBER OF SEQ ID NOS: 168
; SEQ ID NO 10
; LENGTH: 2879
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (611)...(1843)
US-09-906-158-10

```

```

Query Match          17.0%; Score 230; DB 11; Length 2879;
Best Local Similarity 52.7%; Pred. No. 7.7e-52;
Matches 655; Conservative 0; Mismatches 530; Indels 57; Gaps 5;

```

```

Qy      36 TGCCGCTGCTGCTGCCGCTGCTGTGGCTGCTAGTGCTGACGCCTGGCCGGCCGGCCGCCG 95
      ||| ||| | ||| | ||| ||||| | | |||
Db      612 TGCACCTGCAAAGGGCTCTGGTAGTCCTGGCCCTGCTGAACTTGGCCACAATCAGCCTCT 671

Qy      96 GACTGTCCACCTGCAAGACCATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGAGGCCA 155
      ||||| ||| | | ||| | | ||| |||| | | |||
Db      672 CTCTGTCCACTTGCAACACGTTGGACTTCGGCCACATCAAGAAGAAGAGGGTGAAGCCA 731

Qy     156 TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC 215
      || | || |||| | | ||||| |||| ||||| ||| | ||
Db     732 TTAGGGGACAGATCTTGAGCAAGCTCAGGCTCACCAGCCCCCTGAGCCATCGGTGATGA 791

Qy     216 CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG 275
      | | || | ||| | || ||||| ||||| ||||| ||| | |
Db     792 CCCACG-----TCCCTATCAGGTCTCGGCACTTTACAACAGCACCCGGGAGTTGCTGG 845

Qy     276 CCG-----GGGAAAGTGTGCAACCGGAGCCCGAGCCAGAGGCGGACTACT 320
      | | |||| | | | | | ||| | | ||||
Db     846 AAGAGATGCACGGGAGAGGGAGGAAGGCTGCACTCAGGAGACCTCGGAGTCTGAGTACT 905

Qy     321 ACGCCAAGGAGGTCAACCGCTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAAT 380
      | |||| | || | | ||| | | | | | |||
Db     906 ATGCCAAAGAGATCCATAAATTGACATGATCCAGGGAAGTGGCGGAGCAATGAAGTGG 965

Qy     381 TCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGG 440
      | | || | || | ||||| |||| | | |||
Db     966 CCGTCTGCCCCAAAGGAATTACCTCTAAGGTTTTTCGTTTCAATGTGTCCTCAGTGGAGA 1025

Qy     441 TGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCA----- 491
      | | | ||| | | | | | ||| | | |||
Db    1026 AAAATGGAACCAATCTGTTCCGGGCAGAGTTCGGGTCTTGCGGTGCCCAACCCAGCT 1085

Qy     492 -----AGTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATG 539
      || | | |||| | | || | | | | |||
Db    1086 CCAAGCGCACAGAGCAGAGAATTGAGCTCTTCAGATACTTCGACCGGATGAGCACATAG 1145

Qy     540 ATTCCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGACTCACCGGAGTGGCTGT 599
      ||||| | | | | ||| | | ||| | | |||
Db    1146 CCAAGCAGCGCTACATAGGTGGCAAGAATCTGCCACAAGGGGCACCGCTGAATGGCTGT 1205

Qy     600 CCTTTGATGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTT 659

```

Db	1206	CTTTCGATGTCACTGACACTGTGCGCGAGTGGCTGTTGAGGAGAGAGTCCAACCTTGGGTC	1265
Qy	660	TTCGCCTCAGTGCCCACTCTTCCTCTGACA-----GCAAAGATAACACACTCCACG	710
Db	1266	TGGAAATCAGCATCCACTGTCCATGTACACCTTTCAGCCCAATGGAGACATACTGGAAA	1325
Qy	711	TGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGA	770
Db	1326	ATGTTTCATGAGGTGATGGAATCAAATTCAAAGGAGTGGACAATGAAGATGACCATGGCC	1385
Qy	771	ACCGGCCCTTCCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCT	830
Db	1386	GTGGAGACCTGGGGCGTCTCAAGAAGCAAAGGATCACCAACCCACACCTGATCCTCA	1445
Qy	831	CCCGGCACCGCCGAGCCCTGGATACCAACGACTACAAGG-----ATGACGACGACAAGG	884
Db	1446	TGATGATCCCCCACACCGACTGGACAGCCCAGGCCAGGGCAGTCAGAGGAAGAAGAGGG	1505
Qy	885	CCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCT	944
Db	1506	CCCTGGACACCAATTACTGCTTCCGCAACCTGGAGGAGAACTGCTGTGTACGCCCCCTTT	1565
Qy	945	ACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATG	1004
Db	1566	ATATTGACTTCCGCGAGGATCTAGGCTGGAATGGGTCCACGAACCTAAGGGTTACTATG	1625
Qy	1005	CCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGG	1064
Db	1626	CCAACCTTCTGCTCAGGCCCTTGCCCATACCTCCGCGAGCGAGACACAACCCATAGCACGG	1685
Qy	1065	TCCTGGCTCTGTACAACCAGCACAAACCGGGCGCGTCGGCGGCGCCGTGCTGCGTGCCGC	1124
Db	1686	TGCTTGACTATACAACACCCTGAACCCAGAGGCGTCTGCCTCGCCATGCTGCGTCCCCC	1745
Qy	1125	AGGCGCTGGAGCCACTGCCCATCGTGTAACGTGGGCGCGAAGCCCCAAGGTGGAGCAGC	1184
Db	1746	AGGACCTGGAGCCCCTGACCATCTTGTAACATGTGGGCAGAACCCCAAGGTGGAGCAGC	1805
Qy	1185	TGTCCAACATGATCGTGCGTTCCTGCAAGTGCAGCTGAGGCC	1226
Db	1806	TGTCCAACATGGTGGTGAAGTCGTGTAAGTGCAGCTGAGGCC	1847

RESULT 15

US-09-813-271B-7

; Sequence 7, Application US/09813271B

; Patent No. US20020115834A1

; GENERAL INFORMATION:

; APPLICANT:

; (A) Nico Cerletti

; TITLE OF INVENTION: New process for the production of
; biologically active protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. US20020115834A1artis Patent Department

; STREET: 564 Morris Avenue

```

;          CITY: Summit
;          STATE: New Jersey
;          COUNTRY: USA
;          ZIP: 07901
;
;  COMPUTER READABLE FORM:
;          MEDIUM TYPE: Floppy disk
;          COMPUTER: IBM PC compatible
;          OPERATING SYSTEM: PC-DOS/MS-DOS
;          SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;
;  CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/09/813,271B
;          FILING DATE: 20-Mar-2001
;
;  PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: PCT/EP95/02719
;          FILING DATE: 12-Jul-95
;          APPLICATION NUMBER: EPO 94810439.3
;          FILING DATE: 25-Jul-94
;
;  ATTORNEY/AGENT INFORMATION:
;          NAME: Pfeiffer, Hesna J. .
;          REGISTRATION NUMBER: 22640
;          REFERENCE/DOCKET NUMBER: 4-20039C/C1C1/USN
;
;  TELECOMMUNICATION INFORMATION:
;          TELEPHONE: (908) 522-6940
;          TELEFAX: (908) 522-6955
;
;  INFORMATION FOR SEQ ID NO: 7:
;          SEQUENCE CHARACTERISTICS:
;              LENGTH: 336 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: double
;              TOPOLOGY: linear
;
;  MOLECULE TYPE: other nucleic acid
;          DESCRIPTION: /desc = "recombinant hybrid DNA of
;
;  IMMEDIATE SOURCE:
;          CLONE: E. coli LC137/pPLMu.TGF-beta1(44/45)beta3
;
;  FEATURE:
;          NAME/KEY:  mat_peptide
;          LOCATION:  1..132
;          OTHER INFORMATION: /product= "N-terminal 44 amino
;              acids of human TGF-beta1"
;
;  FEATURE:
;          NAME/KEY:  mat_peptide
;          LOCATION:  133..336
;          OTHER INFORMATION: /product= "C-terminal 68 amino
;              acids of human TGF-beta3"
;
;  FEATURE:
;          NAME/KEY:  CDS
;          LOCATION:  1..336
;          OTHER INFORMATION: /product= "hybrid TGF-beta named
;              TGF-beta1-3"
;
;  SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-813-271B-7

```

```

Query Match          16.8%;  Score 227.2;  DB 10;  Length 336;
Best Local Similarity 79.8%;  Pred. No. 2.6e-51;
Matches 268;  Conservative 0;  Mismatches 68;  Indels 0;  Gaps 0;

```

Qy 884 GCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTC 943


```

2:  em_esthum:*
3:  em_estin:*
4:  em_estmu:*
5:  em_estov:*
6:  em_estpl:*
7:  em_estro:*
8:  em_htc:*
9:  gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%		Query				Description
	No.	Score	Match	Length	DB	ID			
	1	715.6	53.0	1072	12	BM562135			BM562135 AGENCOURT
c	2	710.4	52.7	1201	13	BX355682			BX355682 BX355682
c	3	699.4	51.8	983	13	BX335351			BX335351 BX335351
	4	690	51.1	900	13	BX349319			BX349319 BX349319
c	5	658	48.8	1041	13	BX324511			BX324511 BX324511
c	6	619.2	45.9	888	13	BX434425			BX434425 BX434425
	7	619	45.9	1043	12	BM555996			BM555996 AGENCOURT
c	8	612	45.4	950	13	BX383773			BX383773 BX383773
	9	577.4	42.8	859	9	AL530081			AL530081 AL530081
	10	570.6	42.3	713	10	BE312000			BE312000 601154768
	11	562.6	41.7	717	10	BE260971			BE260971 601153715
	12	561.6	41.6	902	13	BQ675698			BQ675698 AGENCOURT
	13	534.4	39.6	925	12	BI818841			BI818841 603037307
c	14	533.8	39.6	841	9	AL530080			AL530080 AL530080
c	15	533.4	39.5	956	12	BI084718			BI084718 602869722
	16	518.6	38.4	1093	12	BQ054305			BQ054305 AGENCOURT
	17	513.8	38.1	871	13	BQ952138			BQ952138 AGENCOURT
	18	510.8	37.9	1049	12	BQ053417			BQ053417 AGENCOURT
	19	506.6	37.6	1013	12	BM469326			BM469326 AGENCOURT

c	20	505.2	37.4	757	14	CA309731	CA309731	UI-H-FT1-
	21	504.8	37.4	943	12	BI909079	BI909079	603070060
	22	502	37.2	821	10	BG748049	BG748049	602705544
	23	501.8	37.2	773	12	BI195242	BI195242	602944518
c	24	498	36.9	1047	13	BX325288	BX325288	BX325288
c	25	478.2	35.4	727	9	AI131171	AI131171	qc15c11.x
c	26	473	35.1	785	9	AI148173	AI148173	qb56d01.x
	27	472.8	35.0	675	10	BE261764	BE261764	601147839
	28	471.8	35.0	637	12	BG938856	BG938856	cn28g07.y
	29	467.8	34.7	748	10	BF682938	BF682938	602117568
	30	464.2	34.4	960	13	BU957001	BU957001	AGENCOURT
c	31	462.6	34.3	722	13	BU632686	BU632686	UI-H-FE1-
c	32	460.8	34.2	928	13	BX412434	BX412434	BX412434
c	33	460.6	34.1	695	14	CA425775	CA425775	UI-H-FE1-
	34	459.6	34.1	659	10	BG500488	BG500488	602544826
	35	458.8	34.0	565	14	CD287835	CD287835	1_M15.abd
c	36	458.6	34.0	697	13	BU633535	BU633535	UI-H-FL1-
c	37	458	34.0	700	13	BU633922	BU633922	UI-H-FL1-
c	38	457.4	33.9	690	13	BU625377	BU625377	UI-H-FG1-
c	39	457.4	33.9	697	14	CA426391	CA426391	UI-H-FE1-
c	40	456.6	33.8	778	9	AI743724	AI743724	wg53b06.x
	41	456.2	33.8	814	12	BI686959	BI686959	603313837
c	42	454	33.7	1015	10	BG683840	BG683840	602651694
	43	453.6	33.6	599	12	BI905963	BI905963	603062849
	44	452.2	33.5	548	12	BM068788	BM068788	ie85a02.y
c	45	451.2	33.4	720	14	CA416525	CA416525	UI-H-FE0-

ALIGNMENTS

RESULT 1

BM562135

LOCUS BM562135 1072 bp mRNA linear EST 20-FEB-2002

DEFINITION AGENCOURT_6562032 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745463
5', mRNA sequence.

ACCESSION BM562135

VERSION BM562135.1 GI:18807966

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1072)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM12768 row: h column: 08

High quality sequence start: 9
High quality sequence stop: 647.

FEATURES Location/Qualifiers
 source 1. .1072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5745463"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_118"
 /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."
BASE COUNT 217 a 356 c 313 g 186 t
ORIGIN

Query Match 53.0%; Score 715.6; DB 12; Length 1072;
Best Local Similarity 82.5%; Pred. No. 2.4e-145;
Matches 886; Conservative 0; Mismatches 144; Indels 44; Gaps 4;

```
Qy      127 GCTGGTGAAGCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCT 186
          |||
Db       20 GCTGGTGAAGCGGAAGCGCATCTAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCT 79

Qy      187 TGCCAGCCCCCGAGCCAGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGGC 246
          |||
Db       80 CGCCAGCCCCCGAGCCAGGGGAGGTGCCGCCCGGCCGCTGCCCCAGGCCGTGCTCGC 139

Qy      247 TCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCC 306
          |||
Db      140 CCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGCC 199

Qy      307 AGAGGCGGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAAGCGGCAACCA 366
          |||
Db      200 TGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCACAACGA 259

Qy      367 AATCTATGATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGA 426
          |||
Db      260 AATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGA 319

Qy      427 GCTCCGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAG 486
          |||
Db      320 GCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAG 379

Qy      487 GCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTG 546
          |||
Db      380 GCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTTCCTG 439

Qy      547 GCGCTACCTCAGCAACCGGCTGCTGGCCCCCAGTGA CTCACCGGAGTGGCTGTCTTTGA 606
          |||
```


Db 440 GCGATACCTCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGA 499
 Qy 607 TGTCACCGGAGTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCT 666
 |||||
 Db 500 TGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCT 559
 Qy 667 CAGTGCCCACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTT 726
 |||||
 Db 560 TAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTT 619
 Qy 727 CAATTCTGGCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCT 786
 |||||
 Db 620 CACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCT 679
 Qy 787 CCTCATGGCCACCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGC 846
 |||||
 Db 680 TCTCATGGCCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGC 739
 Qy 847 CCTGGATACCAACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTT 906
 |||||
 Db 740 -----CCCTGGACACCAACTATTGCTT 761
 Qy 907 CAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCT 966
 |||||
 Db 762 CAGCTCCCGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAGGGACCT 821
 Qy 967 GGGCTGGAAGTGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTG 1026
 |||||
 Db 822 CGGCTGGAAGTGGATCCACGAGCCCCAAGGCTACCATGCCCACTTCTGCCTCGGGCCCTG 881
 Qy 1027 TCCCTACATCTGGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCA 1086
 |||||
 Db 882 CCCCTACATTTGGAGCCTGGACACGCGTACCAGCAAGTCCCTGGCCCTGTACACCCGCCA 941
 Qy 1087 CAACCCGGGCGCGTCGGCGGCGCCGTGCTGC--GTGCCGAGGCGCTGGAGCCACTGCCC 1144
 |||||
 Db 942 TAACCGGGCGCCCTCGGCGGCGGGTTTGTCTGGGCCCCAGGCCCTGGACCCCTGCCC 1001
 Qy 1145 AT--CGTGTA TACGTGGGCCGC--AAGCCCCAAGGTGGAGCAGCTGTCCAACAT 1194
 |||||
 Db 1002 CTTGCGGGTACTACCTGGGGCGCCAAGCCCCAAGTGGGAACACCTGTCCACAT 1055

RESULT 2

BX355682/c

LOCUS BX355682 1201 bp mRNA linear EST 05-MAY-2003

DEFINITION BX355682 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI002YJ05 3-PRIME, mRNA sequence.

ACCESSION BX355682

VERSION BX355682.1 GI:30371987

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1201)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Qy	618	TTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACT	677
Db	688	TTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGA-GGCTTTTCGCCTTAGCGCCCACT	630
Qy	678	CTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCC	737
Db	629	GCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCC	570
Qy	738	GCCGGGGTGACCTGGCCACCATTCACGGCATGAACCGGCCCTTCCTGCTCCTCATGGCCA	797
Db	569	GCCGAGGTGACCTGGCCACCATTCATGGCATGAACCGGCCCTTCCTGCTTCTCATGGCCA	510
Qy	798	CCCCGCTGGAGAGGGGCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCA	857
Db	509	CCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGAC----	454
Qy	858	ACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACGG	917
Db	453	-----ACCAACTATTGCTTCAGCTCCACGG	429
Qy	918	AGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGT	977
Db	428	AGAAGNACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGT	369
Qy	978	GGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCT	1037
Db	368	GGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCTACATTT	309
Qy	1038	GGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCG	1097
Db	308	GGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCG	249
Qy	1098	CGTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACG	1157
Db	248	CCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACG	189
Qy	1158	TGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTTCTGCAAGTGCA	1217
Db	188	TGGGCGGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTTCTGCAAGTGCA	129
Qy	1218	GCTGAGGCCCCGCCCCGCCACAGCCCCGCCACCCGGCAGGCCCGGCCCCACCCCCGCC	1277
Db	128	GCTGAGGTCCCGCCCCGCCCCGCCCGCCCGCAGGCCCGGCCCCACCCCGCCCCGCC	69
Qy	1278	CGCCT-----CACCGGGGCTGTATTTAAGGACATCGTGCCCCAAGCCCACTTGGGAT	1329
Db	68	CCGCTGCCTTGCCCATGGGGGCTGTATTTAA-GACACCCGTCCCAAGCCACNNAMNM	10
Qy	1330	CGATTAAAG	1338
		: : : :	
Db	9	MDATHAAAS	1

RESULT 3

BX335351/c

LOCUS BX335351 983 bp mRNA linear EST 01-MAY-2003

DEFINITION BX335351 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

Qy	518	CTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCC	577
		:	
Db	743	CTGTACCAGAAATACAGCAACAATYCCTGGCGATACCTCAGCAACCGGCTGCTGGCACCC	684
Qy	578	AGTGA CTCA CCGAGTGGCTGTCCTTTGATGTCA CCGGAGTTGTGCGGCAGTGGCTGACC	637
Db	683	AGCGACTCGCCAGAGTGGTTATCTTTTGATGTCA CCGGAGTTGTGCGGCAGTGGTTGAGC	624
Qy	638	CGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTCTTCCTCTGACAGCAAAGAT	697
Db	623	CGTGGAGGGGAAATTGA -GGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGAT	565
Qy	698	AACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCACC	757
Db	564	AACACACTGCAAGTGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCACC	505
Qy	758	ATTCACGGCATGAACCGGCCCTTCTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCAG	817
Db	504	ATTCATGGCATGAACCGGCCCTTCTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCAG	445
Qy	818	CACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGAC	877
Db	444	CATCTGCAAAGCTCCCGGCACCGCCGA-----	418
Qy	878	GACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	937
Db	417	-----GCCCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	364
Qy	938	CAGCTCTACATTGACTTCCCGGAAGGACCTGGGCTGGAAGTGGATTTCATGAACCCAAGGGC	997
Db	363	CAGCTGTACATTGACTTCCCGGAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGC	304
Qy	998	TACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTAC	1057
		:	
Db	303	TACCATGCCAATTTCTGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTGGACACGCAGTAM	244
Qy	1058	AGCAAGGTCCCTGGCTCTGTACAACAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCTGC	1117
		: :	
Db	243	MMCAAGTCTCTGGCCMTGTACAACAGCATAACCC -GGCGCCTCGGCGGCGCCGTNGT-C	186
Qy	1118	GTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGGTG	1177
Db	185	GTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCTGCAAGCCCAAGGTG	126
Qy	1178	GAGCAGCTGTCCAACATGATCGTGCCTTCTGCAAGTGCAGCTGAGGCCCGCCCCGCCCC	1237
Db	125	GAGCAGCTGTCCAACATGATCGTGCCTCCTGCAAGTGCAGCTGAGGTCCCGCCCCGCC-	67
Qy	1238	ACAGCCCCGCCACCCGGCAGGCCCGGGCCCCACCCCCGCCCGCCTACCGGGGCTG	1293
Db	66	----CCGCCCCGCCCGGCAGGCCCGGGCCCCACCCCCGCCCGCCCCCGCTGCCTTG	15

RESULT 4
BX349319
LOCUS

BX349319

900 bp

mRNA

linear

EST 05-MAY-2003

DEFINITION BX349319 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL010YL07 5-PRIME, mRNA sequence.

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1041)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9160.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0AC024DB10NP2&cluster=9160.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0AC024DB10NP2.
 FEATURES Location/Qualifiers
 source 1..1041
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC024YD20"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 176 a 283 c 346 g 219 t 17 others
 ORIGIN
 Query Match 48.8%; Score 658; DB 13; Length 1041;
 Best Local Similarity 81.3%; Pred. No. 8.1e-133;
 Matches 868; Conservative 10; Mismatches 135; Indels 55; Gaps 8;
 Qy 279 GGGAAAGTGTCGAACCGGAGCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCACCC 338
 :||| :| ||||| ||||| ||| :| :| :|||
 Db 1022 SGGGAGAKGSAGAACCGGAGCCCGACCCCTRASCCR--ACTAYACCCCAAGRAGTCACCC 966
 Qy 339 GCGTGCTAATGGTGGAAAGCGGCAACCAAATCTATGATAAATTCAAGGGCACCCCCACA 398
 |||||:|||||:| ||||| ||||| ||||| ||||| ||||| |||||
 Db 965 GCGTGYTAATGKT-GAAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGTACACACA 907
 Qy 399 GCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGC 458
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 906 GCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGTTGC 847
 Qy 459 TCTCTC-GGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAG 517
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 846 TCTCCSSGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAG 787
 Qy 518 CTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCCCC 577

Db	786	CTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCAGCAACCGGCTGCTGGCACCC	727
Qy	578	AGTGACTCACCGGAGTGGCTGTCCTTTGATGTACCGGAGTTGTGCGGCAGTGGCTGACC	637
Db	726	AGCGACTCGCCAGAGTGGTTATCTTTTGTATGTACCGGAGTTGTGCGGCAGTGGTTGAGC	667
Qy	638	CGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTCTTCCTCTGACAGCAAAGAT	697
Db	666	CGTGGAGGGGAAATTGA-GGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGAT	608
Qy	698	AACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGTGACCTGGCCACC	757
Db	607	AACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCACC	548
Qy	758	ATTCACGGCATGAACCGGCCCTTCTCTGCTCCTCATGGCCACCCGCTGGAGAGGGCCAG	817
Db	547	ATTCATGGCATGAACCGGCCCTTCTCTGCTTCTCATGGCCACCCGCTGGAGAGGGCCAG	488
Qy	818	CACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACGAC	877
Db	487	CATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGAC-----	452
Qy	878	GACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	937
Db	451	-----ACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGG	407
Qy	938	CAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTATGAACCCAAGGGC	997
Db	406	CAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGC	347
Qy	998	TACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGTAC	1057
Db	346	TACCATGCCAACTTCTGCCTCGGGCCCTGCCCCCTACATTTGGAGCCTGGACACGCAGTAC	287
Qy	1058	AGCAAGGTCCTGGCTCTGTACAACCAGCACAAACCCGGGCGCGTCGGCGGCGCCGTGCTGC	1117
Db	286	AGCAAGGTCCTGGCCCTGTACAACCAGCATAACCC-GGCGCCTCGGCGGCGCCGTG-TGC	229
Qy	1118	GTGCCGAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTG	1177
Db	228	GTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTG	169
Qy	1178	GAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCAGCTGAGGCCCGCCCCGCC	1237
Db	168	GAGCAGCTGTCCAACATGATCGTGCGCTCTTGCAAGTGCAGCTGAGGTCCCGCCCCGCC	109
Qy	1238	ACAGCCCCGCCACCCGGCAGGCCCGGCCCAACCCCGCCCGCCT-----CACCGGG	1289
Db	108	CGCCCCGCCCGGCAGGCCCGGCCCAACCCCGCCCGCCCGCTGCCTTGCCCATGGGG	49
Qy	1290	GCTGTATTTAAGGACATCGTGCCCCAAGCCCACTTGGGATCGATTAAA	1337
Db	48	GCTGTATTTAAGGACACCCGTCCCCAAGCCCACTGGSCACCCAYTAA	1


```

Qy      576 CCAGTGA CTCA CCGGAGTGGCTGTCCTTTGATGTCA CCGGAGTTGTGCGGCAGTGGCTGA 635
      ||||| ||||| || ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      181 CCAGCGACTCGCCAGAGTGGTTATCTTTTGTATGTCA CCGGAGTTGTGCGGCAGTGGTTGA 240

Qy      636 CCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTCTTCCTCTGACAGCAAAG 695
      ||| ||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      241 GCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGG 300

Qy      696 ATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCA 755
      ||||| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db      301 ATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCA 360

Qy      756 CCATTCA CCGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGGCC 815
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      361 CCATTCA CCGCATGAACCGGCCCTTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGGCC 420

Qy      816 AGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAACGACTACAAGGATGACG 875
      |||| |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      421 AGCATCTGCAAAGCTCCCGGCACCGCCGA----- 449

Qy      876 ACGACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGC 935
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      450 -----GCCCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGC 501

Qy      936 GGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGTGGATTCATGAACCCAAGG 995
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      502 GGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGG 561

Qy      996 GCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCTGGAGCCTAGACACTCAGT 1055
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      562 GCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCTACATTTGGAGCCTGGACACGCACT 621

Qy      1056 ACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCGCGTCGGCGGCGCCGTGCT 1115
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      622 ACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCGCCTCGGCGGCGCCGTGCT 681

Qy      1116 GCGTGCCG CAGGCGCTGGAGCCACTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGG 1175
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      682 GCGTGCCG CAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGCAAGCCCAAGG 741

Qy      1176 TGGAGCAGCTGTCCAACATGATCGTGCCTTCTGCAAGTGCAGCTGAGGCCCCGCCCCGC 1235
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      742 TGGAGCAGCTGTCCAACATGATCGTGCCTCCTGCAGTGCACCTTGAAGGGTCCGGCCCC 801

Qy      1236 CCACAGCCCCGCCACCCGGCAGGCCCGGCCCCACCCCGCCCGCCTCACCGGGG 1290
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      802 CACNCGGNCNNNNNGGNNNNNGGGGGNCCCCGGCGCCGGGCCCCCGCGGCCCG 856

```

RESULT 8

BX383773/c

LOCUS BX383773 950 bp mRNA linear EST 08-MAY-2003

DEFINITION BX383773 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK001YA15 3-PRIME, mRNA sequence.

ACCESSION BX383773

VERSION BX383773.1 GI:30457168

```

KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 950)
AUTHORS       Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE         Full-length cDNA libraries and normalization
JOURNAL       Unpublished
COMMENT       Contact: Genoscope
               Genoscope - Centre National de Sequencage
               BP 191 91006 EVRY cedex - France
               Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
               Library was constructed by Life Technologies, a division of
               Invitrogen. This sequence belongs to sequence cluster 9160.r For
               more information about this cluster, see
               http://www.genoscope.cns.fr/
               cgi-bin/cluster.cgi?seq=CS0DK001AA08NP1&cluster=9160.r. Contact :
               Feng Liang Email : fliang@lifetech.com URL :
               http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
               Faraday Avenue Genoscope sequence ID : CS0DK001AA08NP1.

FEATURES
  source      Location/Qualifiers
               1. .950
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="CS0DK001YA15"
               /cell_type="HELA CELLS COT 25-NORMALIZED"
               /cell_line="HELA"
               /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
               /note="1st strand cDNA was primed with a NotI-oligo(dT)
               primer. Five prime end enriched, double-strand cDNA was
               digested with Not I and cloned into the Not I and EcoR V
               sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT    171 a      260 c      312 g      195 t      12 others
ORIGIN

Query Match      45.4%;  Score 612;  DB 13;  Length 950;
Best Local Similarity  80.2%;  Pred. No. 8e-123;
Matches 789;  Conservative  8;  Mismatches 136;  Indels  51;  Gaps  5;

Qy      353  GAAAGCGGCAACCAAATCTATGATAAATTC AAGGGCACCCCCACAGCTTATATATGCTG 412
          |:||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      943  GRAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGMACACACAGCATATATATGTTC 884

Qy      413  TTCAACACGTCGAGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAG 472
          |||||:| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      883  TTCAAMACATCAGAGCTCCGAGAAGCGGTACCTGVACCCGTGTTGCTCTCCCGGGCAGAG 824

Qy      473  CTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATAC 532
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      823  CTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATAC 764

Qy      533  AGCAATGATTCTTGGCGCTACCTCAGCAACCGGTGCTGGCCCCCAGTGACTCACCGGAG 592
          ||||| || :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      763  AGCAACAATAMCTGGCGATACCTCAGCAACCGGTGCTGGCACCCAGCGACTCGCCAGAG 704

```


cdna clone CS0DD009YM06 5-PRIME, mRNA sequence.

ACCESSION AL530081

VERSION AL530081.2 GI:31067916

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 859)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT On Feb 13, 2001 this sequence version replaced gi:12793574.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9160.r For more information about this cluster, see

<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD009BG03QP1&cluster=9160.r>. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> InVitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DD009BG03QP1.

FEATURES

Location/Qualifiers

source 1..859

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DD009YM06"

/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 179 a 272 c 257 g 148 t 3 others

ORIGIN

Query Match 42.8%; Score 577.4; DB 9; Length 859;

Best Local Similarity 87.0%; Pred. No. 2.6e-115;

Matches 643; Conservative 2; Mismatches 93; Indels 1; Gaps 1;

Qy 6 CCGAGATGGCGCCTTCGGGGCTGCGGCTCTTGCCGCTGCTGCTGCCGCTGCTGTGGCTGC 65

Db 121 CCCCCATGCCGCCCTCCGGGCTGCGGCTGCTGCTGCTGCTGCTACCGCTGCTGTGGCTAC 180

Qy 66 TAGTGCTGACGCCTGGCCGGCCGGCCGGCCGCGGACTGTCCACCTGCAAGACCATCGACATGG 125

Db 181 TGGTGCTGACGCCTGGCCGGCCGGCCGGCCGCGGACTATCCACCTGCAAGACTATCGACATGG 240

Qy 126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185

Db 241 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 300

Qy 186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM195 row: h column: 17
High quality sequence start: 2
High quality sequence stop: 713.

FEATURES
source Location/Qualifiers
1. .713
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3510592"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_19"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 157 a 226 c 211 g 119 t
ORIGIN

Query Match 42.3%; Score 570.6; DB 10; Length 713;
Best Local Similarity 87.5%; Pred. No. 7.4e-114;
Matches 624; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```

Qy      135 AGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCC 194
          |||
Db      1  AGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCC 60

Qy      195 CCCCAGCCAGGGGGACGTGCCGCCCCGGCCCCGCTGCCTGAGGCAGTACTGGCTCTTTACA 254
          |||
Db      61 CCCCAGCCAGGGGGAGGTGCCGCCCCGGCCCCGCTGCCGAGGCCGTGCTCGCCCTGTACA 120

Qy      255 ACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGGAACCGGAGCCCAGCCAGAGGCGG 314
          |||
Db      121 ACAGCACCCGCGACCGGGTGCCGGGGAGAGTGCAGAACCGGAGCCCAGCCTGAGGCCG 180

Qy      315 ACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAAGCGGCAACCAAATCTATG 374
          |||
Db      181 ACTACTACGCCAAGGAGGTCAACCGCGTGCTAATGGTGGAACCCACAACGAAATCTATG 240

Qy      375 ATAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGG 434
          |||
Db      241 ACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAG 300

Qy      435 AAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGT 494
          |||
Db      301 AAGCGGTACCTGAACCCGTGTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGT 360

```

```

Qy      495 TAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCTGGCGCTACC 554
        |||
Db      361 TAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCCTGGCGATACC 420

Qy      555 TCAGCAACCGGCTGCTGGCCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTCACCG 614
        |||
Db      421 TCAGCAACCGGCTGCTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTCACCG 480

Qy      615 GAGTTGTGCGGCAGTGGCTGACCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCC 674
        |||
Db      481 GAGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCC 540

Qy      675 ACTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTG 734
        |||
Db      541 ACTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCG 600

Qy      735 GCCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGG 794
        |||
Db      601 GCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTCCTGCTTCTCATGG 660

Qy      795 CCACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCC 847
        |||
Db      661 CCACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCGAGCC 713

```

RESULT 11

BE260971

LOCUS BE260971 717 bp mRNA linear EST 26-OCT-2000

DEFINITION 601153715F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3509931 5', mRNA sequence.

ACCESSION BE260971

VERSION BE260971.1 GI:9132709

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 717)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM193 row: m column: 04

High quality sequence stop: 713.

FEATURES Location/Qualifiers

source 1..717

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

```

/clone="IMAGE:3509931"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_19"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

```

BASE COUNT      158 a      226 c      213 g      120 t
ORIGIN

```

```

Query Match          41.7%;  Score 562.6;  DB 10;  Length 717;
Best Local Similarity 87.4%;  Pred. No. 4.1e-112;
Matches 627;  Conservative 0;  Mismatches 89;  Indels 1;  Gaps 1;

```

```

Qy      136 GCGGAAGCGCATCGAGGCCATTGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCC 195
          |||
Db      1   GCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCCC 60

Qy      196 CCCGAGCCAGGGGGACGTGCCGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAA 255
          |||
Db      61 CCCGAGCCAGG-GGAGGTGCCGCCCGGCCCGCTGCCCGAGGCCGTGCTCGCCCTGTACAA 119

Qy      256 CAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGCCAGAGGCGGA 315
          |||
Db      120 CAGCACCCGCGACCGGGTGCCCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCCGA 179

Qy      316 CTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACCAAATCTATGA 375
          |||
Db      180 CTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACGAAATCTATGA 239

Qy      376 TAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGGA 435
          ||
Db      240 CAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGA 299

Qy      436 AGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTT 495
          |||
Db      300 AGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTT 359

Qy      496 AAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCT 555
          |||
Db      360 AAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCT 419

Qy      556 CAGCAACCGGCTGCTGGCCCCAGTGACTACCGGAGTGGCTGTCTTTGATGTCAACCGG 615
          |||
Db      420 CAGCAACCGGCTGCTGGCAACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTCAACCGG 479

Qy      616 AGTTGTGCGGCAGTGGCTGACCCGCGAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCA 675
          |||
Db      480 AGTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCCA 539

Qy      676 CTCTTCCTCTGACAGCAAAGATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGG 735
          ||

```

```

Db      540 CTGCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGG 599
Qy      736 CCGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGC 795
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      600 CCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCTTTCCTGCTTCTCATGGC 659
Qy      796 CACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGA 852
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      660 CACCCCGCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCGAGCCCTGGGA 716

```

RESULT 12

BQ675698

LOCUS BQ675698 902 bp mRNA linear EST 15-JUL-2002

DEFINITION AGENCOURT_8036532 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6211917 5', mRNA sequence.

ACCESSION BQ675698

VERSION BQ675698.1 GI:21786532

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 902)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2375 row: c column: 22

High quality sequence stop: 599.

FEATURES

source

Location/Qualifiers

1. .902

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6211917"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_102"

/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 194 a 292 c 255 g 161 t

ORIGIN

Query Match 41.6%; Score 561.6; DB 13; Length 902;
 Best Local Similarity 87.4%; Pred. No. 7.3e-112;
 Matches 615; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy	156	TTCGCGGCCAGATTCTGTCCAAGCTTCGGCTTGCCAGCCCCCGAGCCAGGGGGACGTGC	215
Db	1	TCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCCCCCGAGCCAGGGGGAGGTGC	60
Qy	216	CGCCCGGCCCGCTGCCTGAGGCAGTACTGGCTCTTTACAACAGTACCCGCGACCGGGTAG	275
Db	61	CGCCCGGCCCGCTGCCCCAGGCCGTGCTCGCCCTGTACAACAGCACCCGCGACCGGGTGG	120
Qy	276	CCGGGGAAAAGTGTGCAACCGGAGCCCGAGCCAGAGGCGGACTACTACGCCAAGGAGGTCA	335
Db	121	CCGGGGAGAGTGCAGAACCGGAGCCCGAGCCTGAGGCCGACTACTACGCCAAGGAGGTCA	180
Qy	336	CCCGCGTGCTAATGGTGAAAGCGGCAACCAAATCTATGATAAATTCAAGGGCACCCCCC	395
Db	181	CCCGCGTGCTAATGGTGAAACCCACAACGAAATCTATGACAAGTTCAAGCAGAGTACAC	240
Qy	396	ACAGCTTATATATGCTGTTCAACACGTCGGAGCTCCGGAAGCGGTGCCGGAACCTGTAT	455
Db	241	ACAGCATATATATGTTCTTCAACACATCAGAGCTCCGAGAAGCGGTACCTGAACCCGTGT	300
Qy	456	TGCTCTCTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG	515
Db	301	TGCTCTCCCGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAAGTGGAGCAGCACGTGG	360
Qy	516	AGCTATACCAGAAATACAGCAATGATTCTGGCGCTACCTCAGCAACCGGCTGCTGGCCC	575
Db	361	AGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCAGCAACCGGCTGCTGGCAC	420
Qy	576	CCAGTGACTCACCGAGTGGCTGTCTTTGATGTACCCGAGTTGTGCGGCAGTGGCTGA	635
Db	421	CCAGCGACTCGCCAGAGTGTTATCTTTTGTATGTACCCGAGTTGTGCGGCAGTGGTTGA	480
Qy	636	CCCGCAGAGAGGCTATAGAGGGTTTTTCGCCTCAGTGCCCACTCTTCTCTGACAGCAAAG	695
Db	481	GCCGTGAGGGGAAATTGAGGGCTTTTCGCCTTAGCGCCCACTGCTCCTGTGACAGCAGGG	540
Qy	696	ATAACACACTCCACGTGGAAATTAACGGGTTCAATTCTGGCCGCCGGGGTGACCTGGCCA	755
Db	541	ATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCA	600
Qy	756	CCATTACGGCATGAACCGGCCCTTCTGCTCCTCATGGCCACCCCGCTGGAGAGGGCCC	815
Db	601	CCATTATGGCATGAACCGGCCCTTCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCC	660
Qy	816	AGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCAAC	859
Db	661	AGCATCTGCCAAGCTCCCGGCACCGCCGAGCCCTGGGACACCAC	704

RESULT 13

BI818841

LOCUS

BI818841

925 bp

mRNA

linear

EST 04-OCT-2001


```

QY      126 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATTCGCGGCCAGATTCTGTCCAAGCTTCGGC 185
      |||
Db      266 AGCTGGTGAAGCGGAAGCGCATCGAGGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGC 325

QY      186 TTGCCAGCCCCCGAGCCAGGGGGACGTGCCGCCCGGCCGCTGCCTGAGGCAGTACTGG 245
      |||
Db      326 TCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCCCGGCCGCTGCCCGAGGCCGTGCTCG 385

QY      246 CTCTTTACAACAGTACCCGCGACCGGGTAGCCGGGGAAAGTGTGCAACCGGAGCCCGAGC 305
      |||
Db      386 CCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAGAACCGGAGCCCGAGC 445

QY      306 CAGAGGCGGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAAGCGGCAACC 365
      |||
Db      446 CTGAGGCCGACTACTACGCCAAGGAGGTACCCGCGTGCTAATGGTGGAACCCACAACG 505

QY      366 AAATCTATGATAAAATTCAAGGGCACCCCCACAGCTTATATATGCTGTTCAACACGTCGG 425
      |||
Db      506 AAATCTATGACAAGTTCAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAG 565

QY      426 AGCTCCGGGAAGCGGTGCCGGAACCTGTATTGCTCTCTCGGGCAGAGCTGCGCCTGCTGA 485
      |||
Db      566 AGCTCCGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGA 625

QY      486 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCTT 545
      |||
Db      626 GGCTCAAGTTAAAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTT 685

QY      546 GGCGCTACCTCAGCAACCGGCTGC-TGGCCCCCAGTGACTACCGGAGTGGCTGTCCTTT 604
      |||
Db      686 GGCGATACCTCAGCAACCGGCTGCTTGGCACCCAGCGACTCGCCAGAGTGGTTATCTTTT 745

QY      605 GATGTCACCGAGTTGTGCGGCAGTGGCTGACCCGCAGA--GAGGCTATAGAGGGTTTTT 662
      |||
Db      746 GATGTCACCGAGTTGTGCGGCAGTGGTTGAGCCGTGGACGGGGAACATTGAGGGCTTTT 805

QY      663 GCCTCAGTGCCCACTCTTCCTCTGACAGCAAAGATAAC 700
      |||
Db      806 GCCTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACC 843

```

RESULT 14

AL530080/c

LOCUS AL530080 841 bp mRNA linear EST 23-MAY-2003

DEFINITION AL530080 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD009YM06 3-PRIME, mRNA sequence.

ACCESSION AL530080

VERSION AL530080.2 GI:31067915

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 841)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT On Feb 13, 2001 this sequence version replaced gi:12793573.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9160.r For more information about this cluster, see
[http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD009BG03NP1&cluster=9160.r](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD009BG03NP1&cluster=9160.r). Contact :
Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> InVitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DD009BG03NP1.

BASE COUNT	150 a	240 c	283 g	153 t	15 others
ORIGIN					

Qy	438	CGGTGCCGGAACCTGTATTGCTCT-CTCGGGCAGAGCTGCGCCTGCTGAGGCTCAAGTTA	496
Db	841	CGGTACCTGMACCTTGTTGYTCTCCCCGGGCAGAGCTGCTTCTGCTAAGGCTCAAGTTA	782
Qy	497	AAAGTGGAGCAGCACGTGGAGCTATACCAGAAATACAGCAATGATTCCCTGGCGCTACCTC	556
Db	781	AAAGTGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATWCCTGGCGATACCTC	722
Qy	557	AGCAACCGGCTGCTGGCCCCAGTGACTCACCGGAGTGGCTGTCTTTGATGTCACCGGA	616
Db	721	AGCAACCSGCSGCTGGCACCCAGCSACTCGCCAGAGTGGTTTTCTTTGTGTGTCACCGKA	662
Qy	617	GTTGTGCGGCAGTGGCTGACCCGAGAGAGGCTATAGAGGGTTTTCGCCTCAGTGCCAC	676
Db	661	GTTGTGCGGCAGTGGTTGAGCCGTGGAGGGGAAATTAAGG--TTTCGCCTTAGCGCCNAC	604
Qy	677	TCTTCCTCTGACAGCAAAGATAACA CACTCCACGTGGAAATTAACGGGTTCAATTCTGGC	736
Db	603	GGCTCCTGTAAACAGCAGGGATAACA CACTGCAAGTGAACATCAACGGGTTAACTACCGGC	544
Qy	737	CGCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGCC	796
Db	543	CGCCGAGGTGACCTGGCCACCATTCTTGGCATGAACCGGCCTTCCTGCTTCTCATGGCC	484
Qy	797	ACCCCGCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACC	856

```

      |||
Db      483 ACCCCGCTGGAGAGGGCCCAGCTTCTGCAAAGCTCCCGGCACCGCCGA----- 436
Qy      857 AACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACG 916
      |||
Db      435 -----GCCCTGGACACCAACTATTGCTTCAGCTCCACG 403
Qy      917 GAGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAG 976
      |||
Db      402 GAGAAGAACTGCTGCGTGCGGCAGCTGTACATTAACCTCCGCAAGGACCTCGGCTGGGAG 343
Qy      977 TGGATTTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATC 1036
      |||
Db      342 TGGATCCACGAGCCCAAGGGCTWCCATGCCAACTTCTGCCTCGGGCCCTGCCCCTACATT 283
Qy      1037 TGGAGCCTAGACACTCAGTACAGCAAGGTCTGGCTCTGTACAACCAGCACAACCCGGGC 1096
      |||
Db      282 TGGAGCCTGGACACGCAGGACAGCAAGGTCTGGCCCTGTACAACCAGCATAACCCGG-- 225
Qy      1097 GCGTCGGCGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTAC 1156
      |||
Db      224 GCCTCGGCGGCGCCGTGC-GCGTGCCGCAGGCGCTGGAGCCGCTGCCCTTCGTGTACTAC 166
Qy      1157 GTGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCCTGCAAGTGC 1216
      |||
Db      165 GTGGGCCGCAAGCCCAAGGTGGGGCGGCTGTCCAACATGGTCGTGCGCTCCTGCAAGGGC 106
Qy      1217 AGCTGAGGCCCCGCCCCGCCCCACAGCCCCGCCCCACCGGCAGGCCCCGCCCCACCCCCGC 1276
      |||
Db      105 CGCTGGGGTCCCGCCCCGCCCCGCCCCGCCCCGCGCAGGCCCCGGCCCCCCCCCGCCCCGCC 46
Qy      1277 CCGCCT 1282
      ||: ||
Db      45 CCKGCT 40

```

RESULT 15

BI084718/c

LOCUS BI084718 956 bp mRNA linear EST 20-JUN-2001

DEFINITION 602869722T1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5014421 3', mRNA sequence.

ACCESSION BI084718

VERSION BI084718.1 GI:14503048

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 956)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)


```

Db      595 GCTCCTGTGACAGCAGGGATAACACACTGCAAGTGGACATCAACGGGTTCACTACCGGCC 536
Qy      738 GCCGGGGTGACCTGGCCACCATTACGGCATGAACCGGCCCTTCCTGCTCCTCATGGCCA 797
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      535 GCCGAGGTGACCTGGCCACCATTTCATGGCATGAACCGGCCTTTCCTGCTTCTCATGGCCA 476
Qy      798 CCCCCTGGAGAGGGCCCAGCACCTGCACAGCTCCCGGCACCGCCGAGCCCTGGATACCA 857
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      475 CCCCCTGGAGAGGGCCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTGGAC---- 420
Qy      858 ACGACTACAAGGATGACGACGACAAGGCCCTGGATACCAACTACTGCTTCAGCTCCACGG 917
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      419 -----ACCAACTATTGCTTCAGCTCCACGG 395
Qy      918 AGAAGAACTGCTGCGTGCGGCAGCTCTACATTGACTTCCGGAAGGACCTGGGCTGGAAGT 977
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      394 AGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGT 335
Qy      978 GGATTCATGAACCCAAGGGCTACCATGCCAATTTCTGCCTGGGGCCCTGTCCCTACATCT 1037
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      334 GGATCCACGAGCCCAAGGGCTACCATGCCAATTTCTGCCTCGGGCCCTGCCCCACATTT 275
Qy      1038 GGAGCCTAGACACTCAGTACAGCAAGGTCCTGGCTCTGTACAACCAGCACAACCCGGGCG 1097
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      274 GGAGCCTGGACACGCAGTACAGCAAGGTCCTGGCCCTGTACAACCAGCATAACCCGGGCG 215
Qy      1098 CGTCGGCGGCGCCCGTGCTGCGTGCCGCAGGCGCTGGAGCCACTGCCCATCGTGTACTACG 1157
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      214 CCTCGGCGGCGCCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACG 155
Qy      1158 TGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGTTCTTGCAAGTGCA 1217
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      154 TGGGCCGCAAGCCCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCTTGCAAGTGCA 95
Qy      1218 GCTGAGGCCCCGCCCCGCCACAGCCCCGCCCAACCGGCAGGCCCGGCCCCACCCCGGCC 1277
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      94 GCTGAGGTCCCGCCCCGCC-----CCGCCCGCCCCGGCAGGCCCGGCCCCACCCCGGCC 40
Qy      1278 CGCCTCACCGG 1288
      ||||| |||||
Db      39 CGCCCCCGCTG 29

```

Search completed: October 28, 2003, 00:08:20
Job time : 3352.62 secs